

SOLUCIÓN IOT PARA ANÁLISIS INTELIGENTE EN LA PRODUCCIÓN DE UVA / IOT SOLUTION FOR INTELLIGENT ANALYSIS IN GRAPE PRODUCTION



TRABAJO FIN DE MÁSTER
CURSO 2018-2019

AUTOR
JORGE SÁNCHEZ-BARRIGA PONS

DIRECTOR
GONZALO PAJARES MARTINSANZ

MÁSTER EN INTERNET DE LAS COSAS
FACULTAD DE INFORMÁTICA
UNIVERSIDAD COMPLUTENSE DE MADRID

SOLUCIÓN IOT PARA ANÁLISIS INTELIGENTE EN LA PRODUCCIÓN DE UVA / IOT SOLUTION FOR INTELLIGENT ANALYSIS IN GRAPE PRODUCTION

TRABAJO DE FIN DE MÁSTER EN INGENIERÍA INFORMÁTICA
DEPARTAMENTO DE INGENIERÍA DEL SOFTWARE E
INTELIGENCIA ARTIFICIAL

AUTOR
JORGE SÁNCHEZ-BARRIGA PONS

DIRECTOR
GONZALO PAJARES MARTINSANZ

CONVOCATORIA: JUNIO 2019
CALIFICACIÓN: 8

MÁSTER EN INTERNET DE LAS COSAS
FACULTAD DE INFORMÁTICA
UNIVERSIDAD COMPLUTENSE DE MADRID

1 DE JUNIO DE 2019

AUTORIZACIÓN DE DIFUSIÓN

El/la abajo firmante, matriculado/a en el Máster en Ingeniería Informática de la Facultad de Informática, autoriza a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a su autor el presente Trabajo Fin de Máster: "SOLUCIÓN IOT PARA ANÁLISIS INTELIGENTE EN LA PRODUCCIÓN DE UVA", realizado durante el curso académico 2018-2019 bajo la dirección de Gonzalo Pajares Martinsanz en el Departamento de Ingeniería del Software e Inteligencia Artificial, y a la Biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en Internet y garantizar su preservación y acceso a largo plazo.

Jorge Sánchez-Barriga Pons
1 de Junio de 2019

AGRADECIMIENTOS

A mis padres, hermana, abuelos y mi pareja gracias por el inmenso sacrificio, apoyo, comprensión, cariño, y todo lo que aportáis hacia mi persona, sin vosotros nada de esto sería posible. Gracias a Gonzalo Pajares por ayudarme y apoyarme en el proyecto.

RESUMEN EN CASTELLANO

La productividad en los cultivos agrícolas es cada vez más importante dado que los suelos son menos fértiles y las necesidades de alimentos mayores. La tecnología es clave para ayudar a solventar la falta de alimentos, ya que permite optimizar los recursos y predecir la producción, mejorando así la planificación en la distribución y comercialización, evitando incluso situaciones de riesgo ante la escasez en la producción. El presente trabajo se focaliza en el cultivo de la uva, con necesidades peculiares de cara a la producción, dentro de lo que se conoce como Agricultura de Precisión, donde la automatización es clave. Se plantea una solución conceptual inteligente dentro del paradigma en Internet de las Cosas, IoT (*Internet of Things*).

El planteamiento sobre el desarrollo conceptual en IoT se basa en la captura de datos mediante dispositivos interconectados, para monitorización y análisis de datos. El presente proyecto, como indica su título, se enfoca en estos dos aspectos y especialmente en el análisis inteligente de los datos.

Un dato clave del procesamiento inteligente lo constituye el análisis de imágenes digitales. Sobre dichas imágenes, se realiza a nivel local un procesamiento inteligente, basado en la aplicación de técnicas exclusivas de Redes Neuronales Convolucionales (*CNN, Convolutional Neural Networks*). Mediante la clasificación de las imágenes se obtiene la ratio de uvas-hojas. Este dato, junto con el resto de datos que simulan la captura sensorial, obtenidos de fuentes especializadas, se envían a la nube para su almacenamiento, procesamiento y publicación. El procesamiento inteligente en la nube consiste en definir el correspondiente modelo de regresión lineal múltiple, estimando los parámetros que definen dicho modelo, para poder realizar predicciones con datos actuales. La predicción, además de poder determinar la evolución del cultivo, permite actuar de forma controlada en su caso, tratando de mejorar las condiciones ambientales de cara a variar los parámetros necesarios para mejorar la producción.

PALABRAS CLAVE

Agricultura de Precisión, Viticultura, Internet de las Cosas, IoT, Redes Neuronales Convolucionales, Regresión, Predicción.

ABSTRACT

Productivity in agricultural crops is becoming increasingly important since soils are less fertile with more needs for foods. The technology is key to solve the lack of food, because it allows to optimize resources and to predict the production, improving thus the planning in the distribution and commercialization, avoiding even situations of risk to the shortage in the production. This work is focused on grape cultivation, with peculiar needs towards production, within what is known as precision agriculture, where automation is key. An intelligent conceptual solution is proposed within the Internet of things, IoT paradigm.

The approach on the conceptual development in IoT is based on the capture of data by means of interconnected devices, for monitoring and data analysis. This project, as its title indicates is focused on these two aspects and especially on the intelligent analysis of data.

A key fact for intelligent processing is digital image analysis. On these images, an intelligent processing is carried out at the local level, based on the application of Convolutional Neural Networks (CNN). By classifying the images, the ratio of grapes-leaves is obtained. This data, along with the other data that simulate sensory capture, obtained from specialized sources, are sent to the cloud for storage, processing, and publishing. Intelligent processing in the cloud consists of defining the corresponding multiple linear regression model, estimating the parameters that define the model, to make predictions with current data. The prediction, in addition to being able to determine the evolution of the crop, allows to act in a controlled way when required, trying to improve the environmental conditions in order to control relevant parameters to improve the production.

KEYWORDS

Precision Agriculture, Viticulture, Internet of Things, IoT, Convolutional Neural Networks, Regression, Prediction.

ÍNDICE DE CONTENIDOS

Autorización de Difusión.....	IV
Agradecimientos.....	V
Resumen en castellano	VI
Palabras clave	VII
Abstract	IX
Keywords.....	IX
Capítulo 1 - Introducción	11
1.1 Planteamiento general	11
1.2 Motivación	13
1.3 Objetivos.....	14
1.4 Organización de la memoria.	16
Capítulo 2 - Revisión y descripción de la metodología aplicable.....	17
2.1 El IoT en Viticultura	17
2.2 Aprendizaje automático.....	25
2.2.1 Regresión lineal múltiple.....	27
2.2.2 Redes neuronales convolucionales.....	29
Capítulo 3 - Diseño de la aplicación.....	35
3.1 Arquitectura	35
3.1.1 Módulo tecnológico	35
3.1.2 Módulo lógico.....	38
3.2 Procesamiento de datos.....	40
3.2.1 Análisis de imágenes	40
3.2.2 Predictor	41
Capítulo 4 - Resultados.....	45
4.1 Materiales y recursos.....	45
4.1.1 Imágenes.....	45
4.1.2 Herramientas.....	48
4.2 Procesamiento de las imágenes	48
4.2.1 Entrenamiento de la red	49
4.2.2 Clasificación de imágenes	55
4.3 Resultados del ajuste mediante regresión	61
Capítulo 5 - Conclusiones y trabajo futuro	65
Capítulo 6 - Introduction	69
6.1 General considerations	69
6.2 Motivation	71
6.3 Objectives	72
6.4 Content of this document.....	73
Capítulo 7 - Conclusions and future work.....	75

Capítulo 1 - Introducción

1.1 Planteamiento general

El consumo de alimentos, y el consumo de recursos para su cultivo, es cada día mayor. Naciones Unidas (UN, 2019) en su informe concluye que: *“Cada año, se calcula que un tercio de todos los alimentos producidos, equivalentes a 1300 millones de toneladas por valor de alrededor de 1000 millones de dólares, termina pudriéndose en los contenedores de los consumidores y minoristas, o se estropea debido a las malas prácticas del transporte y la cosecha.”* y *“Si la población mundial llegase a alcanzar los 9600 millones en 2050, se necesitaría el equivalente de casi tres planetas para proporcionar los recursos naturales precisos para mantener el estilo de vida actual”*. Concretamente, el objetivo número 12 del mencionado informe de Naciones Unidas prevé establecer para 2030 medios de producción eficientes que necesiten menos recursos para generar sólo los alimentos necesarios que requiere el planeta. A partir de este informe se deduce que la generación del alimento a escala global supone un 30% de la energía consumida por el planeta y el 22% de producción de gases de efecto invernadero, por lo que se trata de un objetivo clave para el desarrollo de la humanidad.

No solo hay que tener en cuenta el propio alimento en sí, sino también los medios y recursos necesarios para su producción, y uno de los grandes problemas del planeta es la falta de agua dulce no solo para el consumo directo humano, sino también para el animal y el de regadío. Alrededor del 3% del agua del mundo es apta para el consumo animal y/o cultivo, de la cual el 2,5% está congelada en la Antártida, el Ártico y los glaciares. Sólo queda un 0,5% de agua del planeta para sostener un ecosistema que cada día es más frágil por la contaminación de dichas aguas.

En el caso concreto de España el avance de la desertificación resulta preocupante (Gobierno de Aragón, 2018). A las carencias de agua derivadas del alto consumo y falta de conciencia de la sociedad hay que sumar el excesivo uso de agua por cultivos ineficientes. Según el informe de Rodríguez-Casado y col. (2008), que, aunque con más de una década sus conclusiones generales siguen aún vigentes, la huella hidrológica de España se situaba en un volumen anual total cercano a los 48.000 hm^3 , lo que suponía que en España existían 350 embalses con una capacidad para

almacenamiento de un total de 54.000 hm^3 de agua. De este consumo más de la mitad del agua para complacer la demanda procede de otros países. Del total del consumo de agua en España más del 80% se utiliza en la agricultura. Este alto consumo de agua, unido a las extracciones acuíferas ilegales y la sobreexplotación del suelo hacen que éste sea cada vez menos cultivable lo que origina, a su vez, que se necesiten más fertilizantes, más energía y materia prima para la producción, con la consiguiente contaminación que esos químicos conllevan.

En la naturaleza todo está conectado y el nivel actual de consumo de recursos puede llegar a ser insostenible para el planeta Tierra a largo plazo. Para solventar este problema, o al menos mitigarlo en la medida de lo posible, como se ha hecho siempre a lo largo de la historia de distintas maneras, se recurre, y se necesita, la ayuda de nuevas tecnologías, y es aquí donde el Internet de las Cosas, IoT (*Internet of Things*), está llamado a jugar un importante papel. En efecto, gracias al IoT es posible sensorizar e interconectar diferentes dispositivos para realizar un seguimiento de los cultivos y medios de producción a gran escala y a muy bajo coste, permitiendo llevar a cabo estudios detallados sobre lo que está ocurriendo en el desarrollo de los cultivos de cara a la obtención del producto final. Actualmente, sin una sensorización interconectada, resulta complicado determinar exactamente cuál es su situación instantánea, y en tiempo real, así como conocer diversos parámetros que afectan a los mismos, tales como la calidad de la tierra o la humedad entre otros. Y menos aún cómo afectan esos datos a la producción final, para así poder considerar aspectos relevantes de futuro en función de la demanda de los productos, que se traducirá sin duda en un ahorro de costes y optimización de recursos. Con la monitorización de datos reales y actualizados en relación a la situación de una determinada plantación es posible automatizar los medios de producción, con supervisión humana, evolucionando hacia la automatización, para conseguir un rendimiento eficiente, basado en procesos de elaboración eficaz.

La Inteligencia Artificial en conjunción con el paradigma IoT ofrece un amplio abanico de posibilidades y soluciones de cara a la requerida y necesaria automatización, parte de las cuales son precisamente las que se presentan en este trabajo. Dentro de estas soluciones se desarrollan técnicas avanzadas de

procesamiento de datos, tanto a nivel local, y por tanto del lado de la sensórica en lo que se viene a llamar *edge computing*, como a nivel remoto. En este trabajo se ofrecen soluciones de procesamiento en ambos niveles. Por tanto, incidiendo en lo mencionado previamente y considerando el paradigma IoT como la interconexión de dispositivos (sensoriales y de proceso), con trasvase de datos para su monitorización y análisis, el trabajo que se presenta se centra en estos dos últimos aspectos, especialmente en el análisis, para el que se desarrollan y aplican técnicas inteligentes avanzadas.

Así, a nivel local, el núcleo de la automatización se focaliza en el diseño de un método de reconocimiento de imágenes, mediante técnicas de aprendizaje profundo (*Deep Learning*), más concretamente, Redes Neuronales Convolucionales (CNN, *Convolutional Neural Networks*) para obtener valores de una variable clave de análisis, relativa a determinar el grado de desarrollo del cultivo con el fin de incluirla en el sistema de predicción de cara a la producción. La variable mencionada es concretamente la denominada ratio uvas-hojas. Por otro lado, a nivel remoto, esto es en la nube, se compila información relativa a otro tipo de variables (temperatura ambiente, radiación solar, humedad del entorno y tierra, presión atmosférica), como se verá posteriormente. Con todos los datos disponibles se realizan estimaciones con fines de predicción, quedando tanto los datos como los parámetros estimados a disposición y uso público en el correspondiente servicio en la nube, con el fin de poder realizar como predicciones en base a los mencionados parámetros.

1.2 Motivación

El planteamiento del concepto y análisis de la solución que se aborda en el presente trabajo se centra en la monitorización de un viñedo con denominación de origen. El hecho de haber elegido este cultivo y no otro tipo de plantaciones se fundamenta en varias razones, que se sintetizan a continuación. La primera de ellas es que este tipo de cultivos genera un producto nacional con gran impacto dentro de la sociedad española y, además, existe abundante documentación sobre estos cultivos, permitiendo establecer una valoración al respecto. Por otro lado, y como segunda razón, el viñedo en sí necesita unas condiciones climáticas muy concretas, entre las que destaca una pluviometría moderada-baja (400-600 mm como promedio de lluvia al año) o sustratos del suelo específicos para su correcto desarrollo y crecimiento. Además,

y como una tercera razón, un viñedo para ser considerado denominación de origen, como por ejemplo Ribera del Duero, necesita superar una serie de pruebas relativas al grado alcohólico, acidez volátil, sulfuroso total y azúcares reductores, entre otras. Esto se consigue teniendo una producción de uva muy concreta marcada por la normativa de denominación de origen. No se puede exceder el cultivo de uva, pues las propiedades de ésta no cumplirían la normativa vigente. Por otro lado, producciones bajas originarían pérdidas económicas no deseables. Todas estas circunstancias motivan la necesidad de una monitorización en el tiempo, a la vez que se disponen los datos para uso por la comunidad interesada y se realizan los análisis correspondientes para establecer medidas correctoras en su caso, de forma que el IoT proporciona una solución factible a esta problemática.

Un viñedo de denominación de origen como el de Ribera del Duero es un buen objeto de estudio, no solo para poder extrapolar los trabajos realizados a otros cultivos más resistentes como los cereales o el maíz que no necesitan entornos con tanta regularización, sino también, porque necesita tener una cantidad de producción concreta sin excedentes ni escasez, cumpliendo así con el objetivo número doce de Naciones Unidas (UN, 2019).

Desde el punto de vista técnico, la motivación del trabajo queda suficientemente justificada. En efecto, el concepto de desarrollo planteado bajo el paradigma IoT se basa en la captura de datos de procedencia sensorial, que como se ha indicado previamente, se procesan en parte localmente y en parte en la nube. Dado que no existe una implantación real de dispositivos a nivel de campo, la mencionada captura se lleva a cabo, dentro del contexto de IoT, de forma simulada. Así, se dispone de los datos suficientes que permiten determinar las variables objeto de análisis previamente mencionadas. Por tanto, el planteamiento en su conjunto ofrece una solución conceptual y viable en el ámbito del IoT.

1.3 Objetivos

El objetivo general que se propone en el presente trabajo es diseñar un sistema IoT, como solución posible y factible para analizar y optimizar, mediante la aplicación de técnicas inteligentes, un cultivo de viñedo para cumplir el objetivo número doce de Naciones Unidas con el fin de garantizar modalidades de consumo y producción

sostenibles, tal y como se ha indicado previamente. En concreto, se trata de una solución aplicable a un viñedo para producir vino con denominación de origen Ribera del Duero. Dado que no es posible la implantación de las tecnologías sensoriales in situ, la propuesta contempla la captura de datos a partir de repositorios disponibles con carácter público, incluyendo la utilización de imágenes para su procesamiento. De esta forma, se garantiza una solución genérica, de suerte que, para su implantación real, no hay más que sustituir las citadas fuentes por los sensores correspondientes que generan ese mismo tipo de datos.

Dentro de los objetivos específicos y desde el punto de vista global del IoT, se propone desarrollar un sistema inteligente mediante técnicas específicas de aprendizaje profundo, más concretamente del tipo CNN. Esta red procesa las imágenes disponibles en modo local para obtener el porcentaje de hojas de viñedo en relación a la cantidad de uvas presentes (ratio uvas-hojas). Estos datos procesados se suben y publican en remoto (nube), disponiéndose para su visualización pública y procesamiento posterior también en modo remoto, juntamente con el resto de los datos disponibles (temperatura ambiente, radiación solar, humedad del entorno y tierra, presión atmosférica) con fines de predicción. Para la predicción se utilizan técnicas de regresión lineal múltiple. Los datos, a falta de disponer de datos sensoriales reales de un cultivo, se generan de forma aleatoria basados en datos reales de la AEMET (2019) de la estación meteorológica de Aranda de Duero (Burgos).

Por tanto, los objetivos específicos para conseguir la solución IoT integrada, se sintetizan como sigue:

1. Diseño de un procedimiento de obtención de datos.
2. Desarrollo e integración, a nivel local, de un método de procesamiento de imágenes basado en CNN, que incluye sendos módulos de entrenamiento y clasificación
3. Diseño y configuración de la plataforma remota para almacenamiento, procesamiento y publicación de datos en la nube.

4. Análisis de los datos mediante técnicas de regresión multivariable, obteniendo los parámetros correspondientes del modelo, que se publican en la nube.
5. Integración de los módulos que implementan los procesos anteriores.
6. Análisis y valoración de los resultados.

El plan de trabajo previsto tiene como base los objetivos específicos reseñados, contemplando el mismo número de fases que objetivos, y con una planificación equilibrada por semanas a lo largo del tiempo de desarrollo del proyecto.

1.4 Organización de la memoria.

La memoria se estructura en cinco capítulos, claramente diferenciados, donde se detalla la investigación realizada y el desarrollo del proyecto.

En el primer capítulo se detalla el planteamiento general del proyecto, especificando como aspectos generales las razones que motivan la solución IoT propuesta, para resolver el problema de la monitorización de datos en viñedos. Además, se incluyen los objetivos planteados en este trabajo.

En el segundo capítulo se realiza una revisión general sobre aplicaciones de IoT en agricultura, incluyendo la de precisión, en general y en viñedos en particular. También se explican los fundamentos teóricos de las CNN, así como las técnicas de regresión. Son las técnicas de procesamiento de datos utilizadas tanto a nivel local (CNN) como remoto (regresión).

En el tercer capítulo se expone la solución arquitectónica propuesta, incluyendo los módulos local y remoto, incluyendo los mecanismos de comunicación establecidos al respecto, así como la publicación de los datos disponibles en la nube.

En el cuarto capítulo se proporcionan los resultados obtenidos en relación a los distintos módulos desarrollados, así como los relativos al análisis de los procesos realizados.

En el quinto capítulo se incluyen las conclusiones derivadas, así como los posibles trabajos futuros.

Capítulo 2 - Revisión y descripción de la metodología aplicable

A continuación, se realiza una revisión general sobre soluciones automatizadas, con especial énfasis en el ámbito del IoT, en agricultura de precisión a nivel global y en viñedos en particular. Esta revisión establece las bases para abordar la solución propuesta en el presente trabajo.

Se exponen, además, las técnicas inteligentes utilizadas, que comprenden tanto las relativas al modelado y estructura de las CNN dentro del paradigma del aprendizaje automático, como las de regresión, como parte del tratamiento de los datos elaborados. Ambas constituyen la esencia de la solución global inteligente IoT diseñada.

2.1 IoT en Viticultura

El concepto o el término Internet de las Cosas fue descrito por primera vez en 1999 por Ashton (2009) quien especificó la importancia de los datos creados por las propias máquinas. Éstos, a diferencia de los creados por los seres humanos, no tienen la limitación en exactitud, velocidad y cantidad, por ello si permitimos hacer conexiones máquina a máquina (M2M) para intercambiar y generar datos se pueden automatizar tareas de manera casi instantánea sin intervención humana. Esta idea de interacción M2M permite, por ejemplo, la sensorización interconectada de diferentes tipos de dispositivos que proporcionan datos en condiciones varias para su monitorización en tiempo real, así como actuadores automáticos que reaccionan una vez que esa ingente cantidad de datos se ha procesado. También, al generar datos de forma masiva, se pueden llevar a cabo análisis de distinta naturaleza, incluyendo, estudios estadísticos, para realizar pronósticos o predicciones con resultados suficientemente precisos para el ámbito de aplicación de las técnicas utilizadas según el contexto.

En los últimos años la tecnología IoT ha conseguido una robustez suficiente con resultados prometedores en aplicaciones reales, incluyendo diferentes entornos de producción. Algunos ejemplos ilustrativos al respecto son los que se mencionan a continuación, que dan idea de la amplitud de su uso y potencialidad en entornos reales:

- *Gestión de flotas para empresas*, realizando telemetría vía GPS es posible optimizar rutas de forma automática para ser seguidas por vehículos, así como

revisar los tiempos de entrega de productos. Además, sensorizando los propios vehículos se consigue determinar su posicionamiento, calcular posibles averías o prevenir acciones relativas a conducción peligrosa, entre otras posibilidades.

- *Sector alimentario*, con posibilidad de realizar una trazabilidad de los productos de consumo en alimentación para determinar cuál es su estado durante los trayectos en su traslado, detectando así, posibles fallos en su conservación.
- *Protección y seguridad*, ubicando sensores específicos resulta factible detectar posibles fallos en la infraestructura de un edificio, tales como fugas de gas, agua o incluso temblores.
- *Agricultura*, en esta área las diversas soluciones IoT están consiguiendo resultados prometedores, y más específicamente en lo que respecta a lo que se conoce como Agricultura de Precisión (AP), para utilizar y aprovechar los recursos sólo cuando se necesitan y en el momento apropiado. Es justamente en este ámbito donde se focaliza el presente trabajo y más concretamente en viticultura.

La AP es una técnica de cultivo que se apoya en las tecnologías de la información de naturaleza inteligente, para obtener el máximo rendimiento de los cultivos. Por tanto, las actuaciones agrícolas de precisión incluyen procesos inteligentes que se añaden y complementan o sustituyen las tareas esenciales clásicas de la agricultura tradicional con todas las ventajas que ello supone. Son varias las empresas a nivel nacional que han optado por desarrollar tecnologías específicas en el ámbito agrícola dentro de la AP. Por citar sólo dos de ellas, cabe mencionar Innovagri (2019) y Bioagro Technologies (Sánchez, 2015) que vienen desarrollando y aplicando soluciones IoT en AP desde hace ya algunos años.

La primera fase en AP consiste en el estudio y monitorización relativos a la variabilidad del suelo, donde se estudia el trinomio suelo-planta-agua. Para ello se pueden usar dos técnicas bien conocidas: a) *detección remota*, mediante imágenes de satélite o aéreas se estudian las propiedades dinámicas del clima y la evolución vegetal de grandes extensiones; b) *detección a nivel local o sobre el terreno*, se toman muestras específicas del suelo para después analizar sus propiedades. Una vez elegido el suelo donde se va

a cultivar y analizadas sus propiedades, se elige con mucha más precisión las necesidades de aquél para conseguir mejores rendimientos en producción y calidad.

La siguiente fase se centra en la producción, siendo necesario diferenciar aquí distintos niveles de sensorización:

- a. Sensores para la localización y el guiado de la maquinaria agrícola. Así como para posicionar la maquinaria y guiarla durante las tareas propias de los procesos de cultivo (Guerrero y col., 2017). A modo de ejemplo, Navas y col. (2010) concluyen que el uso de GPS y DGPS para el guiado de los vehículos en la siembra o cosecha ahorra entre un 5-7% de combustible.
- b. Sensores en los sistemas eléctricos de los vehículos. Todos los vehículos necesitan ser reparados y a mayor uso mayor es su consumo en combustible y energía, aumentando la cantidad de emisiones de gases nocivos sobre los cultivos. La prevención en la detección de fallos, mediante sensorización permite realizar reparaciones antes de que los problemas de la maquinaria vayan a más, ahorrando costes (Navas y col., 2010).
- c. Sensores de siembra. Principalmente se componen de sensores de distancia basados en ultrasonidos situados en los cuerpos de la maquinaria destinada a tal fin. Estos sensores miden la profundidad de siembra de las semillas, así como la distancia entre las propias semillas, de suerte que su correcta colocación puede aumentar la producción de plantas emergidas de 82 a 102 mm (Navas y col., 2010).
- d. Sensores automáticos de detección de malas hierbas. Entre lo que cabe distinguir a su vez dos tipos, los optoelectrónicos que usan la espectroscopia para distinguir el suelo de la vegetación verde por su reflectividad, y los bi-espectrales que usan imágenes de video, las cuales, convenientemente procesadas mediante diversas técnicas propias de visión por computador permiten diferenciar plantas de malas hierbas de plantas pertenecientes a los cultivos (Burgos-Artizzu y col., 2011; Tellaeche y col., 2011). Los sistemas de esta naturaleza, al detectar una mala hierba, activan actuadores que aplican herbicidas o tratamientos específicos de

forma localizada, esto supone un gran ahorro en los insumos, que suponen alrededor de un 50% en un plazo de 4 a 6 años (Navas y col., 2010).

- e. Sensores de abonado, para determinar las cantidades necesarias en un cultivo. En este sentido se debe determinar cuál es el grado de fertilidad del suelo a abonar mediante diferentes tipos de sensores, eléctricos, electromagnéticos, ópticos, mecánicos y electroquímicos. Estos sensores permiten realizar el abonado en tiempo real y en las áreas requeridas, mientras el vehículo transita entre los cultivos, por ejemplo, en los viñedos.
- f. Sensores de crecimiento. Cuya finalidad es determinar el estado del fruto en relación a su tamaño. Destacan los ópticos con técnicas de procesamiento de imágenes, para detectar entre otras cosas el color, y los de presión para determinar el ancho del fruto, también se pueden usar para obtener propiedades químicas del propio fruto.

Por tanto, la AP requiere de diversos tipos de sensores, los cuales deben estar interconectados para realizar análisis y mapas de rendimiento de los cultivos de la manera más precisa posible. Los datos generados son muy útiles para optimizar el consumo y uso de recursos para reducir costes; así, por ejemplo, si se monitorizan los datos de temperatura del suelo y humedad se puede realizar un estudio de cuán posible es la proliferación de hongos o plagas que afecten al cultivo, reduciendo de forma eficaz el uso de fertilizantes y fungicidas. También sirve para mejorar la calidad del producto, siendo posible medir el tamaño del tallo o el fruto, así como su color para determinar con mayor exactitud la cantidad de agua necesaria. A esto se suma toda la información almacenada o disponible en tiempo real sobre el clima para la realización de informes detallados sobre cómo afectan dichas variables a los cultivos o para activar alertas en caso de superar ciertos umbrales. Además, la conexión de los sistemas informáticos a internet permite acceder a la información de manera completa y sencilla desde cualquier parte o incluso recibir avisos en los sistemas locales, incluyendo los *smartphones*.

En la producción de vino, que es el ámbito donde se ubica el presente proyecto, hay que tener en cuenta problemas específicos que influyen en la calidad, rendimiento y

conservación de la uva (iQConsulting, 2008), además de las mejoras descritas anteriormente en AP.

Las propiedades de la uva que afectan y determinan la posterior calidad del vino, son principalmente: a) *Bajo contenido en azúcar*, de suerte que este problema está relacionado con la mala proporción fruto/hojas, afectando a la graduación del vino, una forma de detectarlo es mediante espectroscopia de infrarrojo cercano; b) *Color*, al igual que en el caso anterior sucede por una mala proporción fruto/hojas o por exceso de fertilización. Sólo se puede detectar mediante técnicas de procesamiento de imágenes digitales; c) *Firmeza de la uva*, una uva débil con falta de consistencia generalmente se debe a una alta acidez, que puede desembocar en enfermedades que destruyen el fruto. Con espectroscopia de infrarrojo cercano y acústica se puede detectar el problema; d) *Quemaduras por el sol*, la falta de agua o de follaje hace que las uvas cambien de color y por tanto sus propiedades. De nuevo con imágenes digitales y sensores de radiación ultravioleta se puede detectar este problema para llevar a cabo acciones preventivas o correctivas; e) *Escobajo negro*, teniendo en cuenta que el escobajo es la parte leñosa que une la uva con el racimo, su coloración oscura indica una mala nutrición por exceso de frío, que se puede detectar con un simple sensor de temperatura; f) *Racimo ralo*, significa que tiene pocas uvas unidas al raquis o parte leñosa central del racimo, lo cual es debido a bajas temperaturas y lluvias intensas. Para paliar este problema se utilizan sensores de temperatura, humedad y precipitación; g) *Pudrición ácida*, esta enfermedad es causada por hongos, levaduras y bacterias por altas temperaturas, un exceso de humedad y racimos muy compactos, la detección se puede llevar a cabo mediante diversos sensores de temperatura, humedad, imágenes digitales y acústicas.

Dentro de los problemas que afectan al rendimiento del viñedo caben distinguir, entre otros: a) *Mal raleo químico*, que aparece cuando los racimos se sitúan muy apretados entre sí o muy sueltos, lo cual favorece el crecimiento de plagas de hongos. Una forma de detectarlo es generando mapas de rendimiento para crear cultivos estructuralmente más uniformes; b) *Daño por heladas*, además de su detección mediante sensores de temperatura, resulta casi imprescindible un sistema de predicción de clima con fines preventivos; c) *Deficiente vigor de la planta*, de forma que una planta

sin suficiente fortaleza origina una baja producción. Esto es debido a diversas causas, pero para una mejor detección es conveniente crear mapas de vigor identificando la cantidad de fruto generado por m^2 , lo que conlleva una mejor redistribución de las viñas; d) *Bajo calibre de uva*, de forma que no se consigue alcanzar un tamaño y peso necesario de las uvas en los racimos. Esto es debido a deficiencias en la nutrición, falta de agua o mala proporción entre fruto/hojas. Existen varios procedimientos para identificar este problema, siendo uno de ellos la generación de mapas de vigor.

Por último, cabe mencionar los problemas que afectan a la conservación de la uva, que afectan también a la calidad del vino, destacan los siguientes: a) *Partiduras finas*, producidas por la interrupción de la cadena de frío, pues las uvas absorben el vapor de agua de las bolsas en las que son almacenadas; b) *Pudrición gris*, producida por el hongo *Botrytis cinerea* resistente al frío, la forma de predecir la aparición de esta infección es analizando la humedad y temperatura del suelo.

Como puede fácilmente deducirse a partir de las consideraciones anteriores, los factores que pueden afectar tanto a la calidad como a la producción de un viñedo son diversos y están conectados entre sí. Por ello, para mejorar y optimizar los recursos de la mejor manera posible se pueden aplicar soluciones ofrecidas por el IoT y aprovechar su potencialidad. Además de las propias soluciones dadas por IoT para la producción de uva hay que tener en cuenta las soluciones en la propia producción del vino, incluyendo su conservación, fabricación, embotellado y transporte.

La compañía Libelium (2019) propone una solución IoT en el ámbito de lo que denomina *Smart Vineyards* (Sancho, 2019). Por parte de la responsable del área de Agricultura, Agua y Medioambiente se plantea un caso de estudio para monitorizar parámetros ambientales decisivos en el desarrollo de los viñedos y consecuentemente en relación a la calidad de la uva y del vino. La temperatura, humedad, radiación solar o el viento son factores determinantes en tal sentido, ya que una adecuada monitorización, permite la realización de análisis exhaustivos con posibilidades de predicción y actuación en su caso. Con tal propósito la solución consiste en desplegar una estación meteorológica convenientemente equipada con dispositivos propios. Esta estación tiene como sensores base los siguientes: a) BME280, que mide parámetros climáticos como la temperatura (entre -45 y $+85^{\circ}\text{C}$), la humedad (entre 0 y 100%) y la

presión atmosférica (entre 30 y 110 KPa); b) LWS, para medir la humedad relativa que absorben las hojas de la viña; c) Watermark-sensor, para medir la humedad de la tierra, este sensor se debe introducir a distintos niveles de profundidad para determinar exactamente cómo afecta a la totalidad de las raíces de las plantas; d) Pt-1000, para medir la temperatura del suelo (entre -50 y 300°C); e) Ecomatik DC3 y DD-S, para medir el diámetro de los troncos y ver cómo es su crecimiento; f) Ecomatik DF para medir el crecimiento de los frutos; g) SQ-110, para medir la radiación solar; h) SU-100, para medir la radiación ultravioleta, muy importante para estimar el color de la uva; i) WS-3000, estación meteorológica equipada con un pluviómetro, un anemómetro para medir la fuerza del viento y una veleta para determinar la dirección del viento; j) TSL2561, sensor de luminosidad. Los datos proporcionados por estos sensores son almacenados en una base de datos en la nube que es fácilmente accesible desde un ordenador o *smartphone* para obtener estadísticas, realizar predicciones o recibir alertas.

Como se aprecia fácilmente, la solución de Libelium es bastante completa, de forma que con pocas unidades sensoriales o estaciones es posible recibir información muy relevante para identificar problemas, tanto en la calidad como en la producción de vino, y proporcionar soluciones apropiadas. Si bien, existen algunos problemas, tales como el bajo calibre de uva, bajo contenido en azúcar y el color que sólo se pueden solventar mediante el procesamiento de imágenes digitales, que permitan obtener el porcentaje relativo existente entre hoja y uva. Es precisamente en este ámbito donde incide de forma relevante el presente proyecto. Mediante el uso de un vehículo equipado con cámaras a ambos lados de éste y un ordenador a bordo con la aplicación que se proporciona en este trabajo, se puede determinar la proporción hoja/uva existente en cada imagen, que puede ser promediada para todo el viñedo. La técnica desarrollada se basa en las CNN, que además se pueden utilizar en el futuro para detección de otras variables con fines de identificación de enfermedades, por ejemplo, sin más que incluir ligeras variaciones en las imágenes utilizadas para el entrenamiento de las CNN. Un prototipo de este vehículo es el que se muestra en la figura 2.1, creado por VineScout (2019). Se trata de un sistema robotizado para la monitorización automática de viñedos. Está equipado con diversas tecnologías sensoriales, entre las que destacan cámaras multispectrales, incluyendo térmicas, para deducir el contenido de nitrógeno de las hojas de las vides, así como otros dispositivos

para determinar el estado del agua de las plantas o la temperatura de las hojas. Por supuesto, posee equipamiento diverso para navegación autónoma. Destacan en este sentido el GPS y los sonares situados en la parte frontal y trasera. El vehículo navega por el viñedo recopilando datos para su posterior análisis con técnicas inteligentes. Posee además comunicación WiFi para transmisión y recepción de información. Este sistema de naturaleza intrínsecamente robótica encaja perfectamente en el paradigma del IoT.

Conviene reseñar en este sentido, que la solución desarrollada en el presente proyecto y basada en la aplicación de las CNN se ha diseñado a nivel conceptual, con el fin de determinar su validez para su implantación en vehículos como el reseñado previamente.



Figura 2.1. Vehículo autónomo (cortesía de F. Rovira-Mas coordinador del proyecto VineScout, 2019).

Una vez obtenido el porcentaje de uva relativo a la cantidad de hojas en los viñedos, junto con los otros datos relevantes, indicados anteriormente, se envían y almacenan en la nube, a través de un servidor remoto, que, como se ha indicado previamente, admite su procesamiento en la propia nube, a la vez que se proporciona su acceso con carácter público. De esta forma se dispone de un sistema IoT con capacidad para analizar la calidad de la uva en un momento determinado o realizar predicciones, todo ello enfocado a determinar la necesidad de aplicar acciones correctoras mediante la aplicación de nutrientes o activación de los sistemas de riego o protección solar, entre otros, a la vez que se establecen criterios para determinar la

aceptación del producto en el mercado y en todo caso optimizar el uso de recursos para obtener la mejor calidad de vino posible.

2.2 Aprendizaje automático

Para entender el significado de aprendizaje automático, conviene realizar antes una revisión sobre el paradigma de Inteligencia Artificial (IA). Se trata de un concepto amplio aplicable a muchas disciplinas, tales como matemáticas, economía, neurociencia, psicología, filosofía, ingeniería informática, lingüística, entre otras, y, en consecuencia, a una amplia variedad de ámbitos, entre los que se encuentra la AP en general y los viñedos en particular.

Existen cuatro tipos de escuelas de pensamiento que la caracterizan. Por un lado, la que define la IA según cómo pueden llegar a pensar las máquinas desde el punto de vista humano (Haugeland, 1985): *"The exciting new effort to make computers think... machines with minds, in the full and literal sense"*, bajo el enfoque de esta escuela es necesario determinar actividades internas del cerebro y cómo piensa un humano. Por otro lado, existe la escuela que aparta el sentido humano y lo reduce a la racionalidad (Charniak and McDermott, 1985): *"The study of mental faculties through the use of computational models"*, el problema básico de esta escuela es que, si bien se pueden codificar de forma lógica muchos pensamientos, existen otros que no lo permiten, como es el caso de la opinión, por ejemplo. También existe la escuela que define la IA según cómo actúan las máquinas, pues considera que una máquina no puede llegar a pensar como tal, considerando la actuación humana (Rich y Knight, 1991): *"The study of how to make computers do things which, at the moment, people are better"*, el test de turing es un buen ejemplo para explicar este tipo de pensamiento, si bien actuar como un humano no siempre es lo más eficiente. Finalmente, también existe la escuela que define la IA por su actuación racional (Poole y col., 1998): *"Computational Intelligence is the study of the design of intelligent agents."*, un agente inteligente (racional), es el que actúa buscando el máximo beneficio. Por lo tanto, se pueden diferenciar dos tipos de IA, la inteligencia artificial fuerte o robusta, que busca alcanzar habilidades cognitivas equiparables a las del ser humano y una inteligencia débil, que mediante algoritmos pretende resolver problemas específicos y no de carácter general, siendo este último enfoque donde se enmarca la aplicación del presente proyecto.

Bajo estos aspectos conceptuales, se puede llegar a comprender el significado de Aprendizaje Automático (AA) o *Machine Learning* (ML). Así, el ML es una rama de la IA que pretende, mediante diferentes algoritmos y técnicas, crear un programa de forma que dado un nuevo ejemplo sea capaz de dar una respuesta lo más satisfactoria posible.

Los algoritmos usados en ML se pueden clasificar en dos grandes grupos (Duda y col., 2001): a) *Aprendizaje supervisado*, donde un instructor muestra al sistema una categoría para cada muestra en el conjunto de datos disponible. El método busca dar resultados esperados para cada muestra, es decir, se tiene una idea de qué valores se deben obtener y busca reducir o minimizar la suma de costes (generalmente una función) durante el entrenamiento con los datos etiquetados; b) *Aprendizaje no supervisado*, no existe un instructor explícito, esto significa que no existe conocimiento previo sobre cómo deben ser los resultados, por lo que se decide u obtiene, mediante técnicas como las de agrupamiento, una estructura de un conjunto de datos sin clasificar.

En el presente trabajo se aplican técnicas de AA supervisado en dos sentidos. Por un lado, en lo que respecta al procesamiento de las imágenes mediante CNN y por otro para usarlo como predictor buscando la relación entre varias variables explicativas y una variable dependiente mediante técnicas de regresión. Un modelo esquemático de AA se proporciona en la figura 2.2. Los datos de entrada, convenientemente depurados, se procesan para obtener los parámetros derivados del aprendizaje durante la fase de entrenamiento, que en el caso de las CNN son los pesos de las capas de convolución y en el caso de los métodos de regresión son los coeficientes asociados con las variables explicativas, todo ello según la correspondiente hipótesis formulada. Con la disponibilidad de nuevos datos, y en consecuencia nuevas variables, se recuperan los parámetros aprendidos para realizar la clasificación o predicción, en su caso. En cualquier caso, se debe establecer una hipótesis de acuerdo a un criterio en función de los datos y el resultado a obtener. En el caso de estudio del presente trabajo los datos son proporcionados por los sensores y se refieren a los especificados previamente, además de las imágenes utilizadas en la CNN.

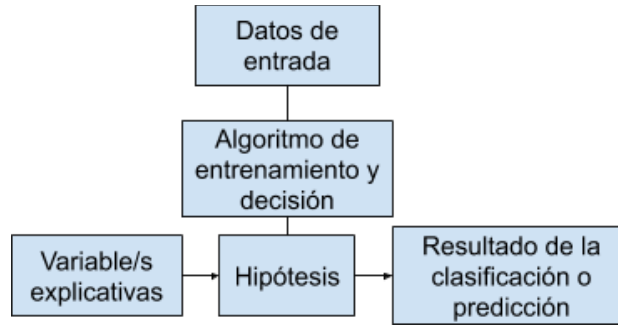


Figura 2.2. Esquema de un sistema de aprendizaje automático.

2.2.1 Regresión lineal múltiple

La regresión lineal ajusta modelos matemáticos lineales para encontrar una relación concreta entre variables independientes o explicativas. El modelo así obtenido se puede usar para poder realizar predicciones.

$$h(x) = \theta_0 + \theta_1 X_1 + \theta_2 X_2 \dots + \theta_n X_n \quad (2.1)$$

asumiendo que $X_0 = 1$, $h(x)$ es la hipótesis, la variable que mide la relación entre las variables explicativas y las variables dependientes. Este modelo se puede expresar en forma vectorial de una forma más sencilla como sigue,

$$X = [X_0, X_1, X_2 \dots X_n] \quad \theta = [\theta_0, \theta_1, \theta_2 \dots \theta_n] \quad h_\theta(x) = \theta^T X \quad (2.2)$$

El método consiste en establecer un vector θ que permita aproximarse lo máximo posible a los valores de las variables dependientes (y). La estimación a esa aproximación se realiza mediante la minimización de una función de coste $J(\theta_0, \theta_1, \theta_2 \dots \theta_n)$ que calcula, en este caso, el error que mide la distancia entre los puntos (x^i, y^i) en el hiperplano $Y = \theta^T X$. Su definición en el caso de la regresión lineal múltiple se establece según la ecuación (2.3):

$$J(\theta_0, \theta_1, \theta_2 \dots \theta_n) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2 \quad (2.3)$$

siendo m la dimensión de los vectores X o y . La clave del éxito consiste en minimizar la función de coste eligiendo diversos θ con un algoritmo determinado. Una manera de encontrar ese mínimo es aplicando la técnica conocida como gradiente descendente. En la figura 2.3 se muestra un gráfico ilustrativo con la evolución de una función de coste y el resultado buscando mínimos con el gradiente descendente.

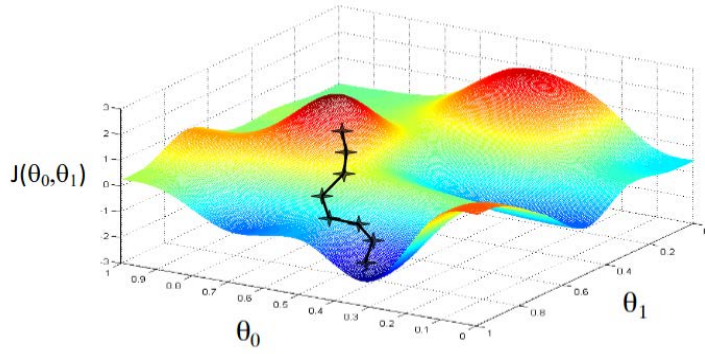


Figura 2.3. Gráfico ilustrando la búsqueda de mínimos mediante el gradiente descendente

La expresión matemática que define la función de gradiente descendente para la regresión lineal múltiple viene dada como sigue,

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (2.4)$$

θ_j establece la relación entre las variables, variando de 1 a n ; α representa la tasa o razón de aprendizaje, esta variable indica la distancia a la que se desplazará el gradiente en cada paso del descenso; un valor muy bajo hace que el proceso sea lento, y uno muy alto podría llegar a saltarse el mínimo buscado.

Otra manera de calcular los θ_j involucrados en la función de coste es mediante la ecuación normal de mínimos cuadrados ordinarios. Este método describe la existencia de un θ_j que establece un mínimo global para θ mediante una medida de ajuste entre la suma de los cuadrados de los residuos, que define la distancia vertical entre un punto y un hiperplano:

$$s(\theta) = \sum_{i=1}^m (y_{(i)} - X_{(i)}\theta) = (y - X\theta)^T (y - X\theta) \quad (2.5)$$

donde $(y - X\theta)^T$ es la matriz transpuesta resultante. Para encontrar el mínimo θ establecido en la suma de cuadrados de los residuos se usa la siguiente ecuación vectorizada,

$$\theta = (X^T X)^{-1} (X^T y) \quad (2.6)$$

donde $(X^T X)^{-1}$ es la matriz inversa. Ambos métodos, ecuación normal y gradiente descendente para encontrar un vector mínimo obtienen iguales resultados. En el caso

planteado en el presente trabajo, se ha comprobado que el menor tiempo de cómputo resulta utilizando la ecuación normal, razón por la que se elige este método.

2.2.2 Redes neuronales convolucionales

Dentro del campo del AA se sitúa el *aprendizaje profundo* (*Deep Learning*, DL), para el cual no se necesita guiar a la máquina en cada una de las fases para que aprenda, sino que la máquina aprende por sí “sola” la estructura de los datos. En esencia no se precisa definir un modelo sino una arquitectura basada en múltiples capas, de suerte que en muchas de esas capas se determinan (aprenden) lo que se conocen como pesos. Se proporcionan ejemplos y es la propia estructura la que determina los valores de dichos pesos en función de las estructuras proporcionadas durante la fase de entrenamiento o aprendizaje.

Gracias al aprendizaje profundo se pueden realizar clasificaciones que antes no eran posibles, el reconocimiento de voz o las señales de tráfico en vehículos autónomos mediante imágenes, son dos buenos ejemplos de ello. Naturalmente, la aplicación planteada en el presente proyecto constituye un ejemplo más de aplicación al respecto.

La clave del DL estriba en la precisión respecto del reconocimiento de patrones que en muchos casos supera a la que un humano puede obtener, como es el reconocimiento de objetos en una imagen. Para crear un modelo correcto se necesita una gran cantidad de datos etiquetados de entrenamiento y una gran capacidad de cómputo, generalmente mediante GPU, ya que es necesario realizar millones de operaciones computacionales con matrices y datos de punto flotante. De hecho, en la actualidad ya se están diseñando las conocidas como TPUs (Tensor Processing Units), que sin duda serán clave en el futuro.

Los modelos de DL funcionan en base a un sistema de etiquetado y de redes neuronales artificiales (RNA) que evalúan los pesos de las conexiones entre las neuronas de las diferentes capas, de forma que a mayor peso mayor interés, de las entradas de las neuronas según el estímulo que se proporcione a las salidas. El término profundo hace referencia a la gran cantidad de capas ocultas, mientras una RNA normalmente tiene dos o tres, una red del tipo DL puede llegar a tener más de 150. Una RNA básica debe

establecer, como cualquier técnica de ML supervisada, una hipótesis de forma que dado un conjunto de entradas se obtenga una salida lo más cercana a una solución proporcionada. La figura 2.4 muestra un esquema de este tipo de redes, que, por cierto, aparecen en las últimas capas del modelo CNN utilizado en el presente trabajo.

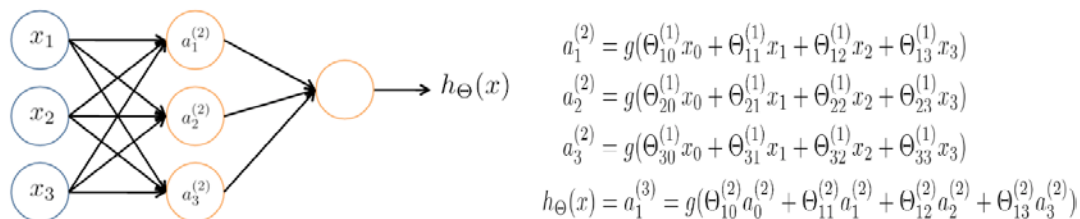


Figura 2.4. Esquema de conexión entre capas y neuronas de una RNA

donde $h_{\theta}(x)$ es la hipótesis planteada; $a_i^{(j)}$ es la unidad de activación i en la capa j ; $\theta^{(j)}$ es la matriz de pesos entre la capa j y la capa $j+1$; g es la función de activación.

Una CNN es un tipo específico de RNA, cuya esencia estriba en lo que se denominan capas de convolución, de ahí su nombre, cuyo fundamento se basa en aprender estructuras a través de los datos suministrados, que simulan el comportamiento de las neuronas de la corteza visual de un cerebro biológico. Por lo tanto, una CNN es una red multicapa de RNA diseñada para reconocer patrones visuales. Una de las CNN más usadas es ALEXNET (ImageNet, 2019; Russakovsky y col., 2015; Krizhevsky y col., 2012; BVLC AlexNet Model, 2019), que ganó la competición ILSVRC-2012 con una tasa de error de un 15.3% frente al segundo competidor que obtuvo un 26.2%. Para comprender este tipo de red primero hay que entender los elementos básicos de su estructura. En la figura 2.5 se muestra las diferentes capas que usa la red ALEXNET.

Se trata de una red con veinticinco capas, previamente pre-entrenada, capaz de clasificar mil objetos diferentes. Los diferentes tamaños de las diferentes capas se establecen mediante las operaciones correspondientes que se especifican seguidamente. En cada capa se aplican operaciones de convolución con núcleos de distintos tamaños especificados mediante la nomenclatura $m \times m$. Por otra parte, los indicadores k , p y s que aparecen en las diferentes capas, hacen referencia respectivamente al número de filtros (con su correspondiente núcleo de convolución) aplicados, al número de ceros añadidos (*padding*) y unidades de desplazamiento del núcleo (*stride*). Como entrada recibe una imagen de dimensión $227 \times 227 \times 3$.

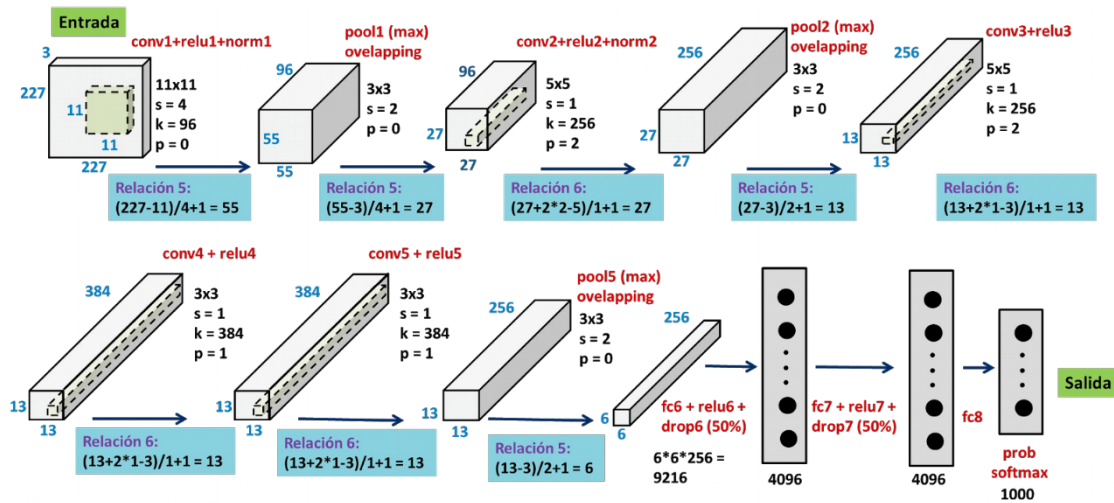


Figura 2.5. Esquema de la red ALEXNET

- **Convolución (conv)**, esta operación matemática combina dos funciones para generar una tercera, básicamente superpone una función sobre otra modificada para obtener una serie de características.

$$S(i, j) = (I * K)(i, j) \quad (2.7)$$

donde el argumento I se conoce como entrada (input), que es un vector o matriz multidimensional. El segundo argumento K , conocido como filtro (kernel) de convolución, es un vector o matriz de parámetros ajustables por el proceso de aprendizaje. Ambas entradas son conocidas como tensores. La salida S se define generalmente como mapa de características (*feature map*).

- **ReLU (relu)**, es una función de activación, de forma que, por cada resultado generado por la convolución, la rectifica linealmente devolviendo 0 para valores negativos y el propio valor de entrada para valores positivos, según se define a continuación, con x siendo la entrada a la neurona.

$$f(x) = \max(0, x) \quad (2.8)$$

- **Normalización (norm)**, la normalización de respuesta local sirve para acelerar la convergencia durante el proceso de aprendizaje. Actualmente se usa la normalización por lotes en lugar de la de respuesta

local. La función de normalización de respuesta local está definida de la siguiente forma:

$$b_{x,y}^i = a_{x,y}^i / (k + \alpha \sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} (a_{x,y}^i)^2)^\beta \quad (2.9)$$

donde N es el número de filtros de la capa dada, $a_{x,y}^i$ es la actividad de la neurona y k, α , n, β son hiperparámetros. En Krizhevsky (2012) se propone $k=2$, $n=5$, $\alpha = 10^{-4}$, $\beta = 0.75$.

- **Pooling (pool)**, la función de pooling modifica la capa de salida agrupando en una porción más pequeña las características, de tal forma que dada una pequeña traslación en la entrada no afecte a los valores de las salidas. Se define como se indica en la ecuación (2.10), teniendo en cuenta que las variables involucradas representan la dimensión de entrada (i), el tamaño del núcleo (*kernel*) de convolución o de *pooling* (k), el desplazamiento del núcleo a lo largo de la imagen (s , *stride*).

$$O = \left\lfloor \frac{i-k}{s} \right\rfloor + 1 \quad (2.10)$$

- **Dropout (drop)**, sirve para no sobresaturar la red neuronal generando soluciones de polinomios de grado demasiado alto. Para ello se propone anular determinado tipo de neuronas de forma aleatoria mediante un hiperparámetro que indique la probabilidad de supervivencia de cada neurona. Las operaciones de este tipo aparecen indicadas como *drop*, correspondiendo en ambos casos al 50% de las neuronas anuladas.
- **Softmax**, se aplica la función exponencial a cada elemento del vector de entrada y se normalizan sus valores por la suma de las exponenciales, lo que asegura que los valores del vector de salida se sitúen en el rango $[0,1]$, proporcionando así la probabilidad de pertenencia a cada una de las clases para una imagen de entrada dada. Su función se define según la siguiente expresión.

$$softmax(x)_i = \frac{\exp(x_i)}{\sum_{j=1}^n \exp(x_j)} \quad (2.11)$$

- **Gradiente descendente**. Técnica de minimización definida previamente en la sección 2.2.1, que se usa también para el ajuste de los pesos de la

red durante el proceso de retropropagación. La función por minimizar se conoce como función *objetivo* o *criterio* y cuando se está minimizando, se le denomina también *cost function*, *loss function* o *error function*.

Como se ha visto, el proceso de entrenamiento de una CNN consiste en extraer determinadas características de las imágenes mediante las operaciones definidas, de forma que se obtienen los valores que definen los núcleos de convolución. Estas características pueden ser bordes, colores, formas u otras estructuras subyacentes en las imágenes. En las primeras capas se usan las funciones de convolución, rectificación lineal para destacar las características importantes y normalización para acelerar la convergencia. Para conectar cada neurona con su capa oculta se utiliza una matriz de pesos, que se modificará según la red neuronal vaya "aprendiendo", y un sesgo, a este par se le conoce como filtro con su correspondiente núcleo. Cada núcleo se genera en función de las características de la imagen a destacar. Seguidamente se usa una capa de pooling para crear una versión condensada de las características derivadas de la información recogida por la capa convolucional, cabe destacar que existirán tantas capas de convolución como de agrupación.

Tras varias etapas de convolución-agrupación se llega a las conexiones de tipo "Fully Connected" o totalmente conectadas, identificadas como *fc*, concretamente en AlexNet se corresponde con la sexta etapa. Esta etapa se caracteriza por realizar la función de conectar todas las neuronas de entrada con todas las neuronas de salida, la función ReLU y la función dropout que permiten reducir el sobreajuste de un modelo, que ocurre por la existencia de diversas capas ocultas con múltiples conexiones, que generan error de tipo aleatorio por el hecho de destacar detalles irrelevantes. Por último, se aplica la función *softmax* definida previamente para obtener la probabilidad para cada una de las clases que queramos analizar.

Para que la red tenga resultados lo más correctos posibles debe ser reentrenada para reajustar los pesos de las capas convolucionales, con respecto a los valores que posee en su modelo original, de forma que las salidas sean las deseadas. En el caso del presente trabajo se han establecido seis clases diferentes de salida (cielo, troncos, tierra, uva-negra, uva-blanca y hojas) por lo que existirá como resultado de la red, después de la operación softmax, un vector cuyos valores corresponden a esas seis clases [y_0 , y_1 ,

y2, y3, y4, y5]. Las seis clases se corresponden con las estructuras relevantes identificadas como tal en las imágenes utilizadas como representativas de los viñedos analizados. En el caso de la red AlexNet los pesos iniciales se heredan del modelo pre-entrenado, dando como resultado una salida, probablemente distante con respecto a la deseada. El objetivo es ajustar (minimizar) la función *de pérdida* durante el entrenamiento, de forma que una salida deseada (ej. [1,0,0,0,0,0]) se corresponda con la que se obtiene para una entrada dada. Este reajuste se realiza mediante retropropagación del error aplicando la técnica del gradiente descendente.

La red debe reentrenarse con el mayor número de ejemplos (imágenes) posibles para conseguir el mejor mínimo posible de la mencionada función de pérdida. Téngase en cuenta en este sentido que una imagen contendrá estructuras diferentes pertenecientes a las distintas clases, lo que redundará en el hecho de utilizar tantas más imágenes para el entrenamiento como sea posible. Por ejemplo, una imagen perteneciente a la clase de uvas negras contendrá también hojas, troncos, o cielo, por ejemplo. No obstante, las características de la red utilizada son capaces de adaptarse a los nuevos ejemplos con un número relativamente pequeño de muestras, lo que le hace ciertamente eficaz para su utilización en el caso objeto de estudio.

Capítulo 3 - Diseño de la aplicación

En el presente capítulo se considera la arquitectura del sistema en su conjunto, que comprende los distintos componentes junto con la lógica de funcionamiento. La aplicación que se propone como solución tecnológica en el ámbito del IoT se aborda desde dos perspectivas estructurales, a saber: arquitectura y procesamiento de datos. Además, se aborda el ciclo completo del tratamiento de datos, desde su captura hasta su procesamiento a distintos niveles y en los diferentes componentes que conforman la arquitectura en su conjunto.

3.1 Arquitectura

La idea de realizar un diseño factible de arquitectura computacional para la monitorización inteligente de cultivos, y más específicamente de viñedos, como es el caso, reside en el planteamiento de la solución a desarrollar bajo el paradigma IoT. En este sentido se propone un diseño arquitectónico del sistema en su conjunto que se desglosa a su vez en dos módulos diferenciados: a) tecnológico y b) lógico.

3.1.1 Módulo tecnológico

Comprende el diseño relativo a los distintos dispositivos y recursos tecnológicos que conforman la propuesta. En este sentido se plantea una solución conceptual, ya que no se dispone actualmente de la tecnología específica para llevar a cabo el despliegue de ésta en campo. Así pues, el problema se aborda considerando la arquitectura tecnológica basada en los diferentes equipamientos disponibles, que se detallan a continuación:

- *Sensores*, encargados de recopilar los datos necesarios para su posterior procesamiento y análisis sobre el estado del cultivo en un momento determinado y para la realización de predicciones a futuro sobre la evolución de ese mismo estado. El planteamiento base consiste en situar en el campo de cultivo una serie de estaciones con los sensores alimentados por baterías cargadas con paneles solares. La identificación de los mismos, en función de los datos que proporcionan para su procesamiento, son los siguientes (Libelium, 2019): a) *BME280*, que mide parámetros climáticos como la temperatura (entre -45 y +85°C), la

humedad (entre 0 y 100%) y la presión atmosférica (entre 30 y 110 KPa); b) *LWS*, para medir la humedad relativa que absorben las hojas de las vides; c) *Watermark-sensor*, para medir la humedad de la tierra, que se introduce a distintos niveles de profundidad con el fin de determinar exactamente el grado de humedad en la totalidad de las raíces de las plantas; d) *Pt-1000*, para medir la temperatura del suelo (entre -50 y 300°C); e) *Ecomatik DC3* y *DD-S*, para medir el diámetro de los troncos y determinar su crecimiento en el tiempo; f) *Ecomatik DF* para medir el crecimiento de los frutos (racimos de uvas); g) *SQ-110*, sensor para medir la radiación solar; h) *SU-100*, para medir la radiación ultravioleta, muy importante para estimar el color de la uva; i) *WS-3000*, estación meteorológica equipada con un pluviómetro, un anemómetro para medir la fuerza del viento y una veleta para determinar la dirección del viento; j) *TSL2561*, sensor de luminosidad. Estos sensores se comunican entre ellos vía *6LoWPAN*, la elección de este estándar reside en que puede asignar una gran cantidad de direcciones IP debido al encabezado IPv6, que es mayor que IPv4. Tiene una topología en estrella o *mesh* perfecta para centralizar la información en un router de borde para posteriormente subir los datos a la nube. Además, tiene tolerancia a fallos como conectividad inestable o el agotamiento de baterías, pudiendo también autoasignar IPs entre los diferentes nodos de la red. Tiene un bajo consumo ya que este protocolo está específicamente diseñado para entornos IoT con tamaños de paquetes reducidos y soporta modo *sleep*. El rango de *6LoWPAN* es adecuado para zonas abiertas, ya que a pesar de que en zonas cerradas no llega a los 15 m, en zonas sin muros se consiguen alcances de hasta 1 Km de distancia, aunque el rango recomendable es de unos 100 metros entre dispositivos, más que suficiente para la monitorización de cultivos de viñedos.

- *Vehículo autónomo*, constituyendo la parte esencial para el reconocimiento de imágenes, así como la detección de problemas en el cultivo de manera precisa. Este vehículo necesita un equipamiento de al menos dos cámaras operando en el rango espectral del visible, con el fin de obtener imágenes en color. Cada cámara se sitúa a ambos lados del

vehículo para capturar imágenes del cultivo mientras transita por una calle del viñedo y sin necesidad de realizar dos pasadas. El vehículo lleva además un computador embebido, encargado de realizar el procesamiento de las imágenes con el fin de determinar el porcentaje de hoja y uva. Dicho procesamiento se lleva a cabo mediante las CNN explicadas en el capítulo precedente. Conviene diferenciar en este sentido, que aquí sólo se ejecutaría la parte de clasificación, mientras que la de entrenamiento, al ser computacionalmente costosa se llevaría a cabo en un procesador con capacidad suficiente para ello, normalmente durante procesos *off-line*, esto es cuando el sistema no esté en acción. En caso necesario, podría disponerse de más sensores para proporcionar datos específicos del terreno, así como para crear mapas de cultivo más precisos; en cualquier caso, todo ello en el ámbito de la AP.

- *Nube*, se dispone de una plataforma que permite recoger y almacenar datos en remoto. Los datos procedentes de las estaciones sensoriales instaladas en el cultivo y los provenientes del robot son almacenados y tratados convenientemente en esta plataforma. Uno de tales tratamientos es la estimación de los parámetros de un modelo de predicción. De esta forma, los datos y los parámetros estimados quedan disponibles con carácter público y con acceso remoto desde distintos dispositivos. En la nube se visualizan representaciones gráficas de los datos recibidos, además de los resultados de predicción respecto de las propiedades del cultivo en general y de la uva en particular.

En resumen, el esquema diseñado para un viñedo inteligente es el que se muestra en la figura 3.1. En el campo de cultivo se sitúan varias estaciones, con los sensores previamente explicados, conectadas entre sí mediante el protocolo 6LoWPAN, existiendo entre ellas una separación máxima de cien metros. Además de estas plataformas pueden existir sensores independientes que se conectan a la red con el mismo protocolo. Todos los datos son enviados al router de borde que tendrá conexión a internet para enviar la información a la nube a través de un procesador central. Por otro lado, el robot podrá conectarse directamente a internet mediante 4G y con este

procesador central. La ventaja de utilizar este protocolo de conexión es que lo hace completamente independiente de su conexión con el router de borde.

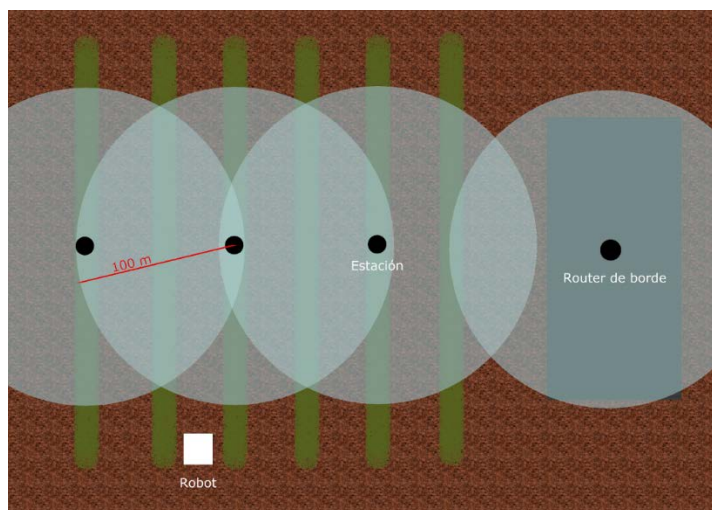


Figura 3.1. Esquema de diseño de sistema inteligente en viñedos

3.1.2 Módulo lógico

El esquema correspondiente a la lógica de diseño se muestra en la figura 3.2. En esta figura se propone un esquema genérico de los distintos componentes que se encuentran interconectados. Según el esquema mostrado en la figura 3.2, la comunicación, y por consiguiente el flujo de datos, entre los distintos elementos que componen el sistema, se establece en doble sentido según se ha descrito en la sección previa 3.1.1, de forma que existe una transferencia de datos entre todos ellos y en todas las direcciones.

- *Conexión sensores y computador local.* Los dispositivos sensoriales se encuentran intercomunicados con el procesador central ubicado en un hipotético Centro de Control (CC) tal y como se plantea en el proyecto RHEA (2014), que aborda el despliegue de una flota de robots con computadores a bordo remotamente interconectados entre sí y con el CC. La conexión con el procesador central permite la recepción de datos procedentes de los sensores según lo establecido. Además, las imágenes recibidas en este computador se procesan mediante las CNN, realizando su correspondiente análisis, mientras que el resto de los datos procedentes de los otros sensores quedan disponibles en el computador local, para su posterior envío a la nube.

- *Conexión computador local y servidor remoto en la nube.* Los datos recibidos en el procesador central son sometidos a dos tipos de procesos. Por un lado, cabe la posibilidad de realizar su procesamiento en dicho computador, de forma que el resultado de éste se envíe a la nube, tal es el caso de las imágenes de los viñedos utilizadas en este trabajo, para obtener la proporción uvas-hojas existente en cada imagen y de forma global en el viñedo. Por otro lado, aquellos otros datos que no necesitan un procesamiento local, o que el mismo sea recomendable en la nube, se envían en crudo a la nube. Éste es el caso para el resto de los datos (temperatura ambiente, radiación solar, humedad del entorno y tierra, nivel de oxígeno de la tierra, presión atmosférica) utilizados en este trabajo.
- *Conexión sensores y servidor remoto en la nube.* Aunque no es la propuesta desarrollada en el presente proyecto, existe la posibilidad de interconectar directamente los dispositivos sensoriales con la nube para el intercambio de datos, y en su caso, la realización de procesamiento a nivel de sensores (*edge computing*), con intercambio de los resultados derivados de su procesamiento.
- *Procesamiento, monitorización y acceso a los datos en la nube.* Con los datos recibidos en la nube se procede por una parte a su almacenamiento, quedando públicamente a disposición de los dispositivos que los requieran, así como para su monitorización y visualización en cualquier dispositivo con acceso a internet. Por otra parte, los resultados obtenidos con los datos procesados, también se disponen públicamente, siendo éste el planteamiento propuesto en este trabajo, tras el correspondiente análisis de regresión. En este tipo de diseños arquitectónicos resulta, en la mayoría de las ocasiones, necesaria la intercomunicación con el CC donde se realizan tareas de monitorización, visualización, planificación o actuación con fines de intervención en los dispositivos remotos o a través de ellos.

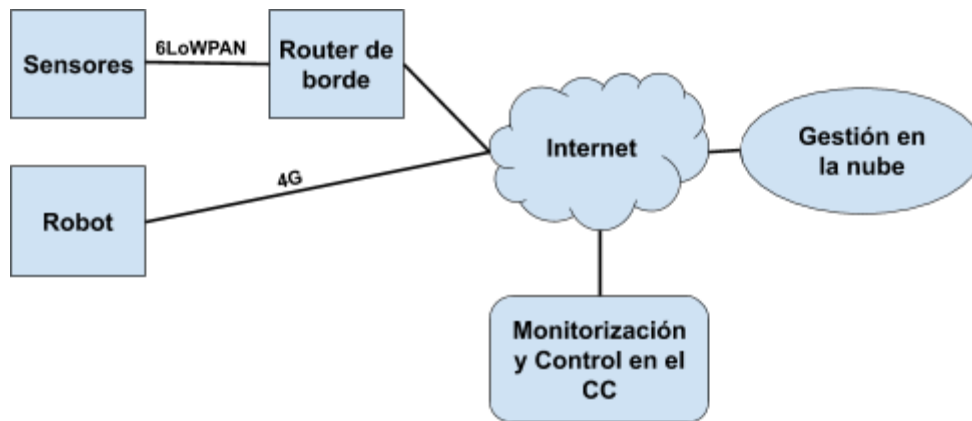


Figura 3.2. Esquema de diseño de sistema inteligente en viñedos

3.2 Procesamiento de datos

3.2.1 Análisis de imágenes

El procesamiento de las imágenes mediante CNN, como es sabido, consta de dos procesos: a) *Entrenamiento*; b) *Clasificación*. La figura 3.3 muestra un ejemplo ilustrativo del procesamiento de imágenes realizado.

Durante la fase de entrenamiento (*off-line*) se proporcionan ejemplos de imágenes, que se corresponden con las cinco clases o categorías utilizadas (cielo, hojas, tierra, uvas-blancas, uvas-negras), que constituyen las clases finales en el modelo propuesto, que se suministran a la red pre-entrenada AlexNet, para proceder a su reentrenamiento. La figura 4.1 muestra cinco ejemplos ilustrativos correspondientes a cada una de las cinco clases. Por tanto, para ello se realiza un reentrenamiento de esta red con los ejemplos disponibles según las clases mencionadas previamente. Las imágenes ejemplo se redimensionan convenientemente hasta conseguir una resolución en tamaño de píxeles de 227x227x3, tal y como requiere el modelo de red mencionado. Los parámetros aprendidos durante este proceso, que en esencia son los correspondientes a los núcleos en las capas de convolución, se almacenan convenientemente en el computador central para su utilización en la siguiente fase de clasificación.

Durante la fase de clasificación (*on-line*), se recuperan dichos parámetros, que se proporcionan a la arquitectura de la misma red AlexNet. Dada una imagen, ésta se particiona convenientemente en sub-imágenes, lo más próximas al tamaño requerido

por la red (227x227x3), con el fin de realizar el mínimo proceso de redimensión necesario. Cada una de estas sub-imágenes se proporciona a la red, obteniendo así un porcentaje de elementos pertenecientes a cada clase. Es decir, para cada una de estas sub-imágenes se obtiene el correspondiente porcentaje de cielo, hojas, tierra, uvas-blancas y uvas-negras. Finalmente, con los porcentajes de uvas (blancas o negras) y el de hojas se calcula la proporción uvas-hojas considerando el conjunto de sub-imágenes utilizadas. De esta forma, se obtiene el porcentaje global uvas-hojas para la imagen completa de entrada, que es una de las variables utilizadas para el análisis predictivo realizado en la nube. El cómputo se puede realizar bien a nivel de imagen o de viñedo global, una vez que se obtiene la media del conjunto de imágenes utilizadas para todo el viñedo. En el presente trabajo la opción utilizada es a nivel de imagen.

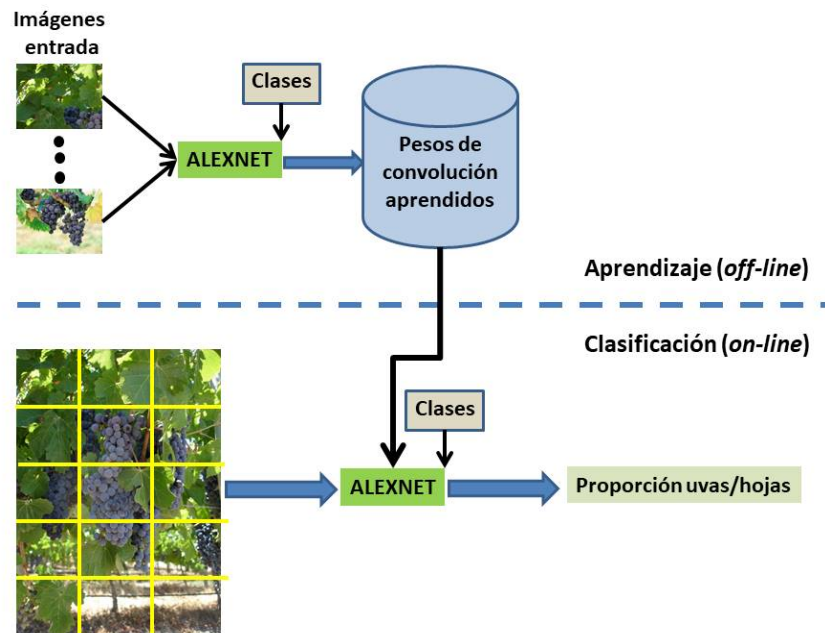


Figura 3.3. Procesamiento de imágenes

3.2.2 Predictor

Como toda solución de naturaleza predictiva en el ámbito del aprendizaje automático, el modelo debe ser definido y sus parámetros estimados, tras lo cual se pueden realizar previsiones con la llegada de nuevos datos en base a los datos disponibles. En este caso se aplica la técnica de regresión lineal para la predicción de las propiedades de la uva en base al estado del cultivo en cuestión.

Por las razones expuestas previamente, en el sentido de que no se dispone de un despliegue sensorial en campo, en la fase de estimación de los parámetros se crean dos ficheros con siete mil datos generados de forma aleatoria. El fichero de datos de entrada contiene los datos relativos a los siguientes estados del cultivo: precipitación, radiación, temperatura atmosférica, humedad atmosférica, presión atmosférica, temperatura del suelo, humedad del suelo y el ratio uva/hoja. Por otro lado, los datos de salida, que son los datos dependientes de los de entrada se corresponden con las siguientes propiedades de la uva: graduación de Baume, Ph, acidez, peso del racimo y antocianos (Rivera del Duero, 2018). Como se ha indicado previamente, todos ellos se generan de forma aleatoria, tomando los máximos y mínimos como referencia, en base a los datos proporcionados por la Agencia Estatal de Meteorología (AEMET, 2013), cuya representación se puede observar en la figura 3.4, que muestra una distribución de las temperaturas medias anuales de Castilla y León, y por los Índices de Maduración de octubre de 2018 del Consejo Regulador de Ribera del Duero (PIM Ribera Duero, 2018). Todos estos datos, por tanto, son generados en el procesador central, simulando así la llegada de datos procedentes de los sensores, siendo enviados a la nube. Conviene reseñar en este sentido, una diferencia procedimental respecto a lo que sería una operativa normal en el análisis temporal. En efecto, los datos almacenados en la nube llevan asociados una etiqueta temporal que determina el momento de tiempo, con expresión de año, día, mes, hora, minuto y segundo. Por tanto, durante la generación aleatoria de los datos de las variables se les asigna la correspondiente etiqueta temporal, también simulada y secuencial en el tiempo. De esta forma, los datos quedan convenientemente almacenados en la nube. Con estos datos se procede a la estimación de los parámetros del modelo de regresión, según se ha explicado en la sección 2.2.1. Estos mismos parámetros quedan también desplegados y almacenados en la nube, cuya implementación, como se indica posteriormente, se realiza en la plataforma ThingSpeak (2019) que da soporte a la arquitectura propuesta según los diseños mostrados en las figuras 3.1 y 3.2. La disponibilidad pública de los datos y parámetros estimados en la nube tiene la gran ventaja de que desde cualquier dispositivo con acceso remoto a ellos se pueden realizar visualizaciones, análisis de interpretación y predicciones sobre la evolución de los cultivos y su estado.

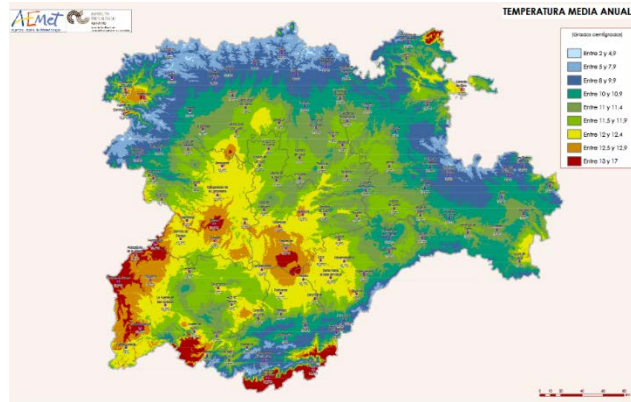


Figura 3.4. Temperaturas medias anuales de Castilla y León
(Fuente Agencia Estatal de Meteorología, AEMET)

Capítulo 4 - Resultados

En el presente capítulo se realiza el correspondiente análisis de resultados, que comprende la descripción de los materiales y recursos utilizados, incluyendo las imágenes y las herramientas de procesamiento y análisis de datos.

Se realiza un análisis exhaustivo de los resultados relativo al procesamiento de las imágenes mediante las CNN en sus fases de entrenamiento y clasificación. El análisis se extiende a los resultados obtenidos mediante regresión, con fines de predicción.

4.1 Materiales y recursos

4.1.1 Imágenes

El conjunto de imágenes utilizadas durante los experimentos es de 200 imágenes (40 por cada una de las cinco clases) para el entrenamiento de la red CNN. Cada imagen contiene diferentes elementos correspondientes a las diferentes clases, de forma que en la misma imagen pueden aparecer tanto uvas (blancas o negras) como hojas, cielo o tierra en distintas proporciones. A continuación, se describe de forma genérica las imágenes correspondientes a cada una de las clases.

Como se ha indicado previamente, para el entrenamiento de la red CNN con el fin de detectar uvas blancas se han utilizado 40 imágenes. En la figura 4.1 se muestran cuatro imágenes representativas procedentes de los repositorios PIXNIO (2019a, b), Semillon grapes (Wikimedia 2019f) y Pixabay (2019). Todas ellas catalogadas en las categorías CC (Creative Commons) y CC0 (Creative Common Zero) de dominio público universal. Sobre ellas se realizan recortes de los racimos conteniendo las uvas objetivo esencial del entrenamiento.



Figura 4.1. Imágenes de entrenamiento para la clase uvas blancas

Para el entrenamiento de la red con muestras de cielo se han utilizado otras 40 imágenes, similares a los ejemplos mostrados en la figura 4.2 correspondientes a los repositorios Iberieli Wineyards (Wikimedia 2019g), Barossa Valley 14 (Wikimedia 2019h), Galician vineyard with wide vine spacing (Wikimedia 2019i), Growing-grapes (Wikimedia 2019j), Tannat_Aout_1 (2019), Montepulciano_d_Abruzzo_04_(RaBoe) (Wikimedia 2019b) y Trte (Wikimedia 2019d). Todas ellas catalogadas en las categorías CC (Creative Commons) y Free Art Licence y por tanto de dominio público.



Figura 4.2. Imágenes de entrenamiento para la clase cielo

Para la clase tierra se han utilizado también 40 imágenes similares a las mostradas en la figura 4.3 con carácter ilustrativo, procedentes de los repositorios Montepulciano_d_Abruzzo_04_(RaBoe) (Wikimedia 2019b), Tannat_Aout_1 (Wikimedia 2019c), Trte (Wikimedia 2019d), Cepa Vieja de Mencía (Wikimedia 2019e) y Barossa Valley plantation (Wikimedia 2019f). Todas ellas catalogadas en las categorías CC (Creative Commons) y Free Art Licence.



Figura 4.3. Imágenes de entrenamiento para la clase tierra

El entrenamiento para la clase uvas negras se ha llevado a cabo igualmente con 40 imágenes, similares a las mostradas en la figura 4.4, procedentes de los repositorios

marius-ciocirlan-414588-unsplash (Unplash 2019a), Samuel-zeller-339493-unsplash (Unplash 2019b), uva-negra (PIXNIO 2019c), vinedo-mixto2 (PIXNIO 2019d) y Australia_Red_Grapes (Wikimedia 2019k). Todas ellas catalogadas en las categorías CC (Creative Commons), CC0 (Creative Common Zero) y Free Art Licence de dominio público.



Figura 4.4. Imágenes de entrenamiento para la clase uvas negras (serie I)

Finalmente también, para el entrenamiento de la clase uvas negras se han utilizado 40 imágenes similares a los ejemplos ilustrativos mostrados en la figura 4.5, obtenidas de los repositorios Australia_Red_Grapes (Wikimedia 2019k), Iberieli Wineryards (Wikimedia 2019g), Samuel-zeller-339493-unsplash (Unplash 2019b), Semillon grapes (Wikimedia 2019l), Tannat_Aout_1 (Wikimedia 2019c) y Growing-grapes (Wikimedia 2019j). Todas ellas catalogadas en las categorías CC (Creative Commons), CC0 (Creative Common Zero) y Free Art Licence con acceso público.



Figura 4.5. Imágenes de entrenamiento para la clase uvas negras (serie II)

4.1.2 Herramientas

Se ha utilizado Matlab (MathWorks, 2019), así como los toolboxes de Deep Learning, Image Processing y Computer Vision con toda su potencialidad, junto con la plataforma ThingSpeak (2019) soportada por Mathworks con sus propiedades y recursos. ThingSpeak es una plataforma específicamente diseñada para operativa IoT, que permite recoger, almacenar, procesar y analizar datos de distinta naturaleza, entre los que se encuentran los procedentes de los dispositivos sensoriales bajo el diseño planteado en este trabajo. Estos datos pueden ser procesados en la nube mediante herramientas Matlab pudiendo ejecutar los programas correspondientes. Los datos se muestran en gráficas según valor y tiempo mediante canales públicos o privados. ThingSpeak proporciona APIs y protocolos de comunicación para acceso a los datos en modo escritura y lectura. Además, permite generar acciones de control cuando se producen ciertos eventos, con fines de actuación sobre dispositivos, cuando éstos dispongan de interfaces al respecto. Entre algunas de las acciones mencionadas se incluye el envío de mensaje vía Twitter, la ejecución de código Matlab o el envío de una petición HTTP. Por ejemplo, si en un momento determinado se sobrepasan ciertos niveles en los valores de las variables relativas al cultivo o la predicción prevé datos alarmantes se puede enviar un aviso por los procedimientos descritos, con el fin de activar algún mecanismo preventivo. La plataforma ThingSpeak proporciona, por tanto, una solución adecuada para el diseño conceptual planteado en el presente trabajo, que permite, como se ha indicado, tanto el almacenamiento de los datos como la estimación de los parámetros del modelo de regresión con fines de predicción del estado del cultivo de uva, junto con su visualización remota.

4.2 Procesamiento de las imágenes

Tal y como se describe en la sección 3.2.1 el procesamiento de las imágenes consta de sendos procesos de *entrenamiento* y *clasificación*. Bajo este planteamiento, a continuación, se proporcionan detalles significativos de los resultados en relación a los mismos.

4.2.1 Entrenamiento de la red

Como se ha indicado previamente, el proceso de entrenamiento se lleva a cabo con el modelo de red AlexNet, para la que se ha configurado la capa de salida de forma que contenga las cinco clases utilizadas en este trabajo en lugar de las mil que contiene la configuración original pre-entrenada.

El número de imágenes utilizadas para cada clase se proporcionan en la tabla 4.1, junto con un ejemplo ilustrativo para cada clase. Obsérvese, cómo para cada clase se han utilizado las 40 imágenes indicadas previamente, con resultados satisfactorios, aunque siempre mejorables, como se verá posteriormente. En este sentido conviene poner de manifiesto la potencialidad de este tipo de redes y la capacidad de ingerir características generales, como las procedentes de las imágenes con las que ha sido entrenada, que son útiles en aplicaciones con un reducido número de ejemplos de entrenamiento. Hay que tener en cuenta que la red AlexNet se ha entrenado con la base ImageNet que contiene del orden de 1,2 millones de imágenes de 1000 clases diferentes. Por tanto, las 200 imágenes utilizadas en el presente trabajo representan un porcentaje mínimo de muestras, con también un número reducido de clases.






Clase	Ejemplo representativo	Número de imágenes
Cielo		40
Tierra		40
Hojas		40
Uvas-negras		40
Uvas-blancas		40

Tabla 4.1 Clases de entrenamiento con ejemplos ilustrativos de cada clase

Como puede observarse, y se ha mencionado previamente, en cada uno de los ejemplos mostrados, las imágenes correspondientes a cada clase contienen mayoritariamente el elemento que representan, pudiendo contener a su vez otros elementos que aparecen mayoritariamente en el resto de las clases, tal es el caso de las uvas (blancas o negras) donde aparecen uvas preferentemente frente a hojas que también están presentes. Este hecho aporta una contribución importante propia de las redes CNN, de tal forma que no es necesario preocuparse por seleccionar específicamente los contenidos de cada imagen respecto de cada clase, tan sólo es necesario preocuparse de que en esa selección se encuentre una representación de la clase correspondiente.

La figura 4.6 muestra la evolución de una fase de entrenamiento (*off-line*) con indicación de la fecha de realización y el tiempo empleado. Se observa, la información del procesamiento bajo un computador con Intel Core i7 2.0 GHz (4th generation) con 8 GB de RAM y Windows 10 como sistema operativo de 64-bits. Se trata de un tiempo suficiente para el propósito que se pretende, teniendo en cuenta que el entrenamiento no es crítico desde el punto de vista operativo en este tipo de sistemas.

Los parámetros y configuraciones establecidos en relación al proceso de entrenamiento son los que se describen a continuación:

1. *Muestras de entrenamiento y validación.* Indican respectivamente, como su nombre indica, el número de muestras utilizadas para el entrenamiento y la validación, obtenidas a partir del total de muestras disponible. En este trabajo, del total de 200 muestras disponibles, el 70% (140) se utilizan para el entrenamiento y el 30% para la validación (60).
2. *Epochs e iteraciones.* Representa el paso de todas las muestras disponibles para el entrenamiento y por tanto las 140 reseñadas previamente. Se han establecido 4 *epochs* con 11 iteraciones por *epoch*, resultando un total de 44 iteraciones.
3. *Precisión en la validación (validation accuracy).* Precisión en el conjunto de imágenes utilizadas para validación, esto es en las 60 imágenes disponibles para este propósito.

4. *Dimensión por lotes (mini-batch size)*. Define la cantidad de imágenes que se utilizan en cada iteración.
5. *Razón de aprendizaje (learning rate)*. Determina la rapidez con la que la red aprende según el método de actualización de los parámetros que constituyen el objetivo del aprendizaje.
6. *Patience*. Indica la detención del proceso tras el número de *epochs*, fijado por dicho parámetro, para el que no se consiguen mejoras relevantes en la precisión.

La tabla 4.2 define los valores de estos parámetros para el proceso de aprendizaje llevado a cabo en uno de los experimentos realizados, que son suficientemente representativos respecto del conjunto de pruebas de entrenamiento llevadas a cabo.

Parámetros de aprendizaje	Valores
Muestras de entrenamiento	140 (70%)
Muestras para validación	60 (30%)
Método de optimización	gradiente descendente
Razón de aprendizaje inicial	1e-4
Disminución de la razón de aprendizaje	0.2 cada 5 epochs
Número máximo de epochs	4
Dimensión por lotes	10

Tabla 4.2 Parámetros de aprendizaje y valores

En la figura 4.6 se muestra el progreso relativo al proceso de entrenamiento, observándose en cada *epoch* e iteración los resultados relativos a la precisión y al ajuste relativo a la función de coste o *loss function*.

En la figura 4.7 se muestra el resultado final del entrenamiento, en este caso se ha obtenido una precisión en la validación del 100%, que puede considerarse ciertamente exitoso, teniendo en cuenta que el aprendizaje se ha realizado en un computador bajo una CPU y por tanto sin GPU. El tiempo total utilizado ha sido de 2 minutos y 34 segundos, que entra dentro de los límites razonables para un proceso *off-line* de tal naturaleza. Obsérvese cómo a partir de la primera *epoch* los resultados correspondientes a la

validación y la precisión en relación al ajuste son aceptables, en el sentido de que se aproximan ya hacia el 100% finalmente conseguido.

De estos resultados se deduce que tanto el diseño y modelo de red utilizado como el número de muestras empleadas son adecuados en el contexto de los experimentos llevados a cabo.

Se ha experimentado con otros modelos de redes, a saber: VGG-16 y VGG-19 (2019) siguiendo las pautas establecidas en Russakovsky y col., (2015) y Simonyan y Zisserman (2014), así como con GoogleLeNet (2019) y ResNet-18 (2019). No se ha llegado a conclusiones relevantes sobre la mejora de resultados con estos modelos respecto de AlexNet, ni en precisión ni en tiempos de cómputo, a pesar de que GoogLeNet contiene configuraciones de capas de tipo *inception* específicamente dedicadas a mejorar los tiempos de cómputo. Por esta razón, se ha decidido realizar todos los experimentos con la mencionada red AlexNet.



Figura 4.6. Progreso del proceso de entrenamiento

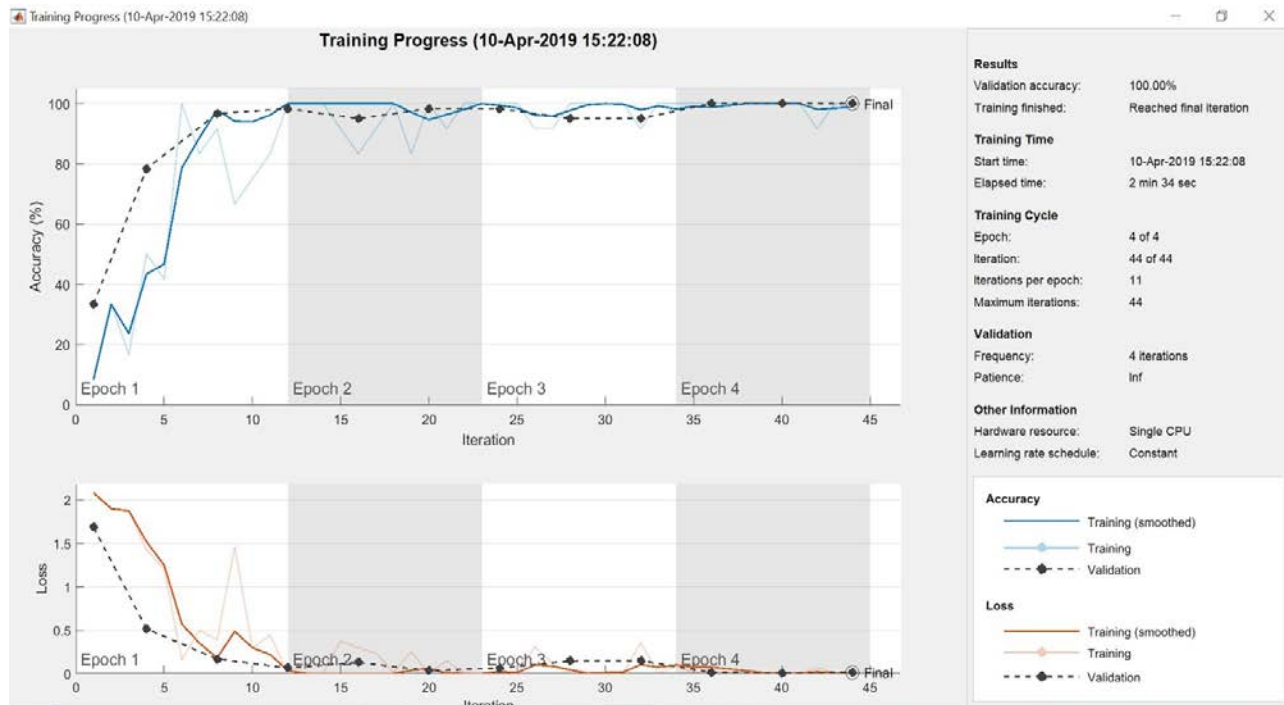


Figura 4.7. Resumen final del proceso de entrenamiento

La figura 4.8 muestra los resultados obtenidos tras el aprendizaje en la primera capa de convolución (*conv1* en la figura 2.5) para una arquitectura de red del tipo AlexNet. Se muestran los 96 filtros correspondiente a esta capa, según se ha definido en la mencionada figura 2.5. Cada filtro posee tres canales, tal y como también se especifica en dicha figura, de ahí su visualización en color. Esta representación coincide de forma muy aproximada con la representación de los conocidos filtros o funciones de Gabor (Daugman, 1980, 1985; Goodfellow y col., 2016). De esta forma, como era de esperar, se establece la conocida similitud con los modelos neuro-computacionales establecidos como base de la justificación de este tipo de redes y su alto rendimiento a la hora de describir las características de los objetos de cara a su reconocimiento.

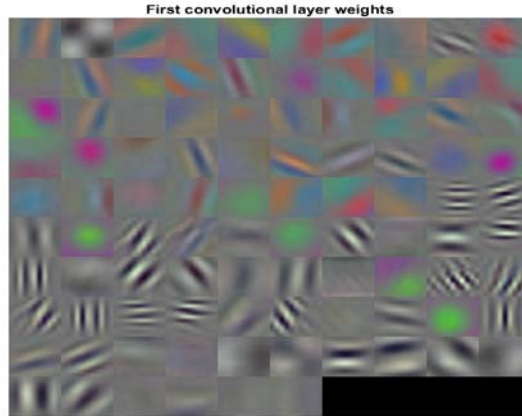


Figura 4.8. Representación de los pesos aprendidos en la primera capa de convolución (conv1)

A modo de ejemplo, también con carácter ilustrativo, en la tabla 4.3 se muestran tres núcleos de convolución para la capa correspondiente identificada como *conv5* en la figura 2.5, que posee 384 núcleos (K) de dimensión 3×3 . Estos filtros no tienen, en principio una interpretación clara relacionada con aspectos neurobiológicos. En cualquier caso, poseen la capacidad de generar a este nivel lo que se conoce como mapas de características, que contribuyen eficientemente a la caracterización de la imagen de entrada con fines de extraer propiedades orientadas a la clasificación de su contenido, como es el caso del reconocimiento de las estructuras presentes en las imágenes y en especial las partes vegetales para caracterizar la proporción de uvas frente a hojas, que es el objetivo clave.

Núcleos	Valores
$K(:, :, 3, 1)$	-0.0124 -0.0065 0.0073 -0.0331 -0.0168 0.0055 -0.0240 -0.0018 0.0141
$K(:, :, 192, 1)$	-0.0285 -0.0179 -0.0177 -0.0159 -0.0193 -0.0110 -0.0141 -0.0201 -0.0105
$K(:, :, 192, 256)$	-0.0161 -0.0194 0.0080 0.0015 -0.0235 -0.0242 0.0047 0.0142 -0.0108

Tabla 4.3. Pesos aprendidos en la capa de convolución cinco (conv5)

4.2.2 Clasificación de imágenes

Una vez realizado el proceso de entrenamiento, se está en disposición de proceder con la clasificación, esto es el proceso *on-line* o decisión, para una imagen de entrada. Para ello se sigue el esquema propuesto en la sección 3.3.1 y esquematizado en la figura 3.3. Así, dada una imagen de entrada a la red AlexNet, ésta progresa hacia adelante, pasando por las distintas capas de la red, obviamente con la misma configuración del modelo previamente entrenado, si bien ahora con los pesos aprendidos durante el entrenamiento y disponibles en las capas de convolución.

A modo de ejemplo, dada la imagen de entrada a la red AlexNet mostrada en la figura 4.9, su paso por la primera capa de convolución (*conv1* en la figura 2.5) genera las 96 imágenes filtradas resultantes mostradas en la figura 4.10. Se observa en esta figura cómo no todos los filtros producen el mismo efecto. Los más activos son aquellos cuyas intensidades y detalles aparecen más marcados, frente a aquellos otros cuyo nivel de activación es considerablemente menor y a veces irrelevante, como ocurre con los resultados donde no aparecen detalles o éstos son mínimos. En este sentido y de cara a mejorar el rendimiento desde el punto de vista de coste computacional, a veces se opta por elegir los núcleos más activos. En el caso del ejemplo indicado, éste se corresponde con el resultado mostrado en la imagen de la figura 4.11, que proporciona el máximo nivel de activación en términos absolutos respecto del resto de canales. En cualquier caso, obsérvese, cómo en esta primera capa se detectan los detalles de la imagen de entrada. Si analizamos las capas más profundas, por ejemplo, la capa de convolución *conv5* de la figura 3.3, los detalles quedan más difuminados en favor de la identificación de otras características inherentes a la imagen de entrada, tal como las componentes espectrales, en este caso el color. Esto aparece en la imagen de la figura 4.12 donde se muestran los resultados generados por los 256 filtros que componen esta capa.



Figura 4.9. Imagen de entrada a la red AlexNet

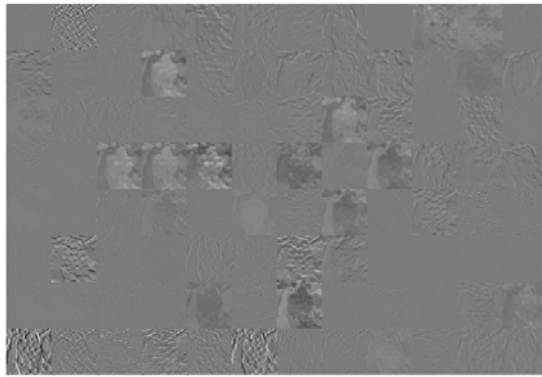


Figura 4.10. Resultados del filtrado con los 96 núcleos de la capa conv1



Figura 4.11. Resultado del canal con el máximo nivel de activación en la capa conv1

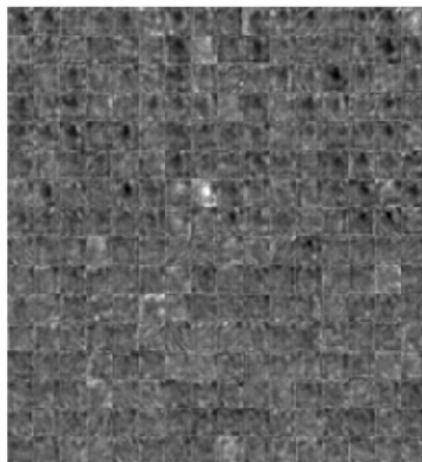


Figura 4.12. Resultados del filtrado con los 256 núcleos de la capa conv5

Una vez entrenada la red se usan varias imágenes de ejemplo para determinar el grado de precisión adquirido en relación al aprendizaje realizado. Tal y como se describe en la sección 3.3.1 y se refleja en la figura 3.3 cada imagen de entrada se trocea en cuadros de dimensión $W \times W$ píxeles, de forma que dependiendo de sus dimensiones se obtiene un número diferente de cuadros. Cada uno de estos cuadros define un recorte de la imagen original, los cuales constituyen así cada una de las entradas a la red. Para cada uno de dichos cuadros se obtiene el porcentaje de estructuras pertenecientes a cada una de las cinco clases (cielo, suelo, hojas, uvas blancas y uvas-negras), con las que se entrenó la red. Una vez procesada la imagen completa, se determina el porcentaje promedio para las mencionadas estructuras teniendo en cuenta el número de cuadros utilizados sobre la correspondiente imagen. Estos porcentajes promedio son los que se muestran y analizan para cada una de las imágenes que aparecen a continuación y que constituyen una base representativa del conjunto total de imágenes utilizadas en los diferentes experimentos.

Como se ha indicado previamente, cada uno de esos cuadros pasa a través de las distintas capas que conforman el modelo AlexNet hasta llegar a la capa *softmax*, figura 2.5. Tal y como se describe en la sección 2.2.2, la salida de esta capa proporciona las probabilidades de pertenencia de la imagen de entrada a cada una de las clases con las que se ha reentrenado la red, que son en definitiva los porcentajes que intervienen en el cómputo del porcentaje global promediado para la imagen completa.

Conviene señalar que para cuadros de dimensión $W = 20$ píxeles los tiempos empleados por la red para realizar la clasificación se sitúan en promedio en torno a 0.07s (70 ms). Teniendo en cuenta que el procesamiento se realiza en Matlab, bajo sus características de lenguaje interpretado, se trata de tiempos aceptables. En cualquier caso, bajo sistemas operativos y lenguajes de tiempo real esos tiempos se reducen considerablemente. Este extremo quedó claramente puesto de manifiesto durante el proyecto RHEA (2014) donde se procesaban cuadros de imagen bajo similares consideraciones computacionales, verificando que estos procesamientos permitían la navegación de los robots a razón de 4 a 6 Km/h, que son velocidades aceptables bajo el paradigma y consideración de la AP.

En la figura 4.13 se muestra una imagen original junto con los porcentajes promediados relativos a las clases con $W = 200$. Se observa visualmente una correspondencia muy aceptable en lo que respecta al contenido de estructuras de clase en la imagen y los mencionados porcentajes, ello a pesar de que el suelo con hierba no lo detecta como suelo propiamente dicho sino como hojas. Esto se debe al hecho de que la red no está entrenada para el reconocimiento de este tipo de estructuras, es decir hierbas. Bien es cierto, que no se realiza un análisis comparativo exhaustivo para determinar los porcentajes de aciertos y fallos con técnicas de medida precisas. Esto es así, dado que el presente trabajo no está enfocado al análisis y desempeño de la red CNN sino a la solución IoT propuesta.

Tanto en este caso como en los siguientes que se analizan, los porcentajes se expresan en términos de probabilidad con rangos entre 0 y 1, de forma que dichos valores se corresponden con porcentajes del 0% y 100% respectivamente. Los porcentajes promedio para cada clase a partir de la imagen mostrada en la figura 4.13 son: 0.3738 de cielo, 0.4963 de hojas, 0.1212 de tierra, 0.0038 de uvas blancas, 0.0048 de uvas negras.

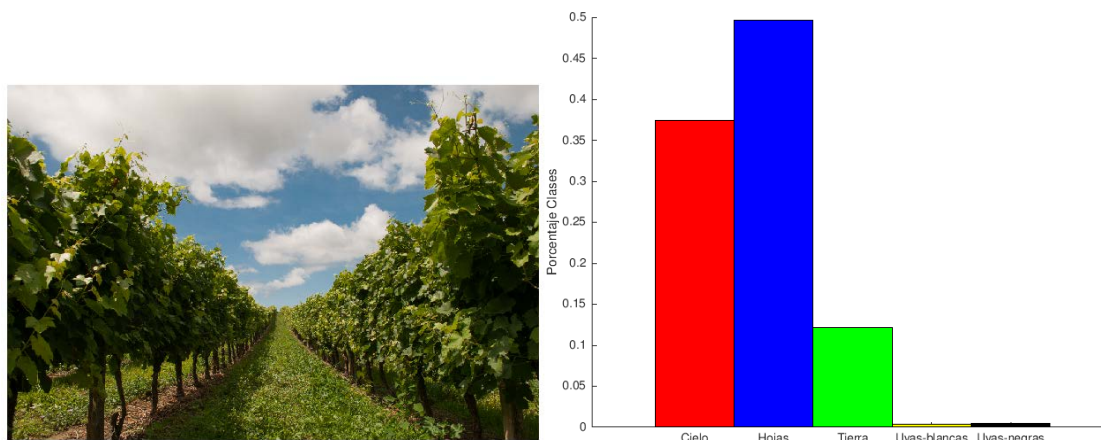


Figura 4.13. Resultados de clasificación para la imagen Growing-grapes.jpg

En la figura 4.14, los resultados se sitúan más próximos en relación al contenido de la imagen, ya que se observa la existencia de un predominio de uvas negras y hojas frente al resto, aunque los porcentajes no sean elevados en términos absolutos. Los porcentajes promedio para cada clase son: 0.0037 de cielo, 0.3450 de hojas, 0.1709 de tierra, 0.0879 de uvas blancas, 0.3924 de uvas negras.

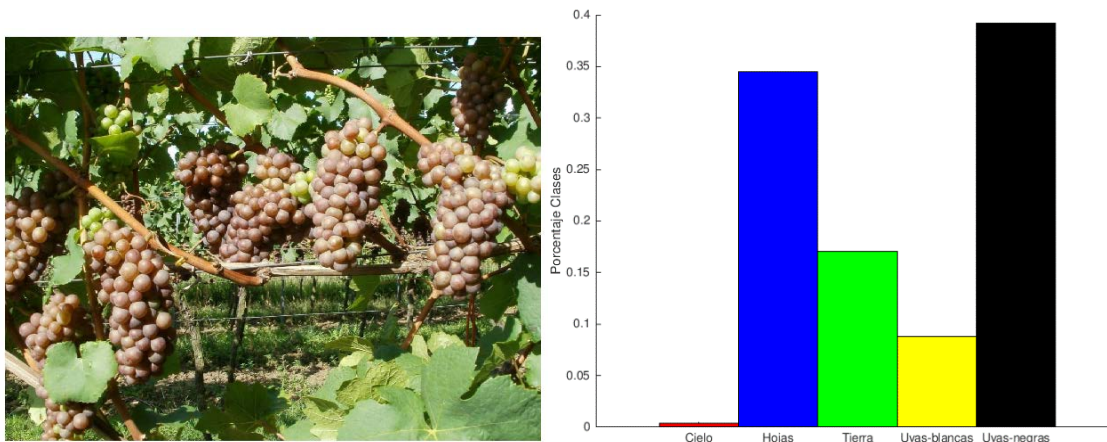


Figura 4.14. Resultados de clasificación para la imagen viñedo-negras.jpg

En la figura 4.15, se observa que prácticamente toda la imagen se compone principalmente de cielo y tierra, con predominio de esta última categoría. Los resultados de la clasificación son coherentes en relación al contenido de la imagen y las respectivas clases. Los porcentajes promedio para cada clase son: 0.4008 de cielo, 0.1326 de hojas, 0.4513 de tierra, 0.0004 de uvas blancas, 0.0149 de uvas negras. Se observa un predominio, en términos relativos, de las clases que aparecen mayoritariamente visibles, observándose además cómo el porcentaje de uvas negras está relacionado con la apreciación cualitativa y el contenido en la imagen de este tipo de estructuras, es decir de uvas negras.

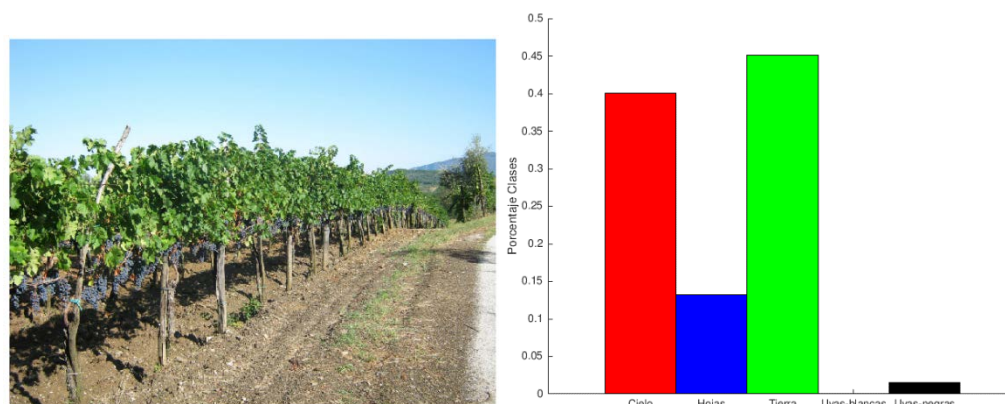


Figura 4.15. Resultados de clasificación para la imagen Trte.jpg

En relación a la figura 4.16 se observa que los resultados no son apreciablemente tan precisos como en los dos casos anteriores, obteniéndose unos porcentajes promedio para cada clase de 0.4600 de cielo, 0.2236 de hojas, 0.2999 de tierra, 0.0003 de uvas

blancas, 0.0162 de uvas negras, lo que discrepa un tanto de lo que cualitativamente se aprecia en la realidad sobre la imagen procesada. Esto se debe al tamaño de los cuadros en los que se divide la imagen ($W \times W$) tal y como se ha explicado previamente. En este caso concreto, la imagen 1200px-Barrosa_Valley_18 posee una resolución espacial en píxeles (tamaño) mucho mayor que las anteriores, por lo que con un recuadro de dimensión $W = 200$ el número de recuadros en los que se divide la imagen es mayor, de forma que teniendo en cuenta que la red CNN está entrenada con imágenes, en su mayoría, con las uvas relativamente próximas respecto del campo de vista de la cámara, hace que en este caso concreto, al situarse éstas alejadas, no haya ejemplos significativos con los que se ha entrenado la red. En estos casos además de considerar la posibilidad de reentrenar la red para este tipo de situaciones, la solución para obtener mejores resultados viene por el incremento del tamaño de la ventana que define el recuadro, tal y como ocurre con un tamaño de $W = 50$, observándose ahora cómo el porcentaje de uvas negras se ve incrementado, pasando a ser del 0.1542 frente al 0.0162 que se obtuvo previamente.

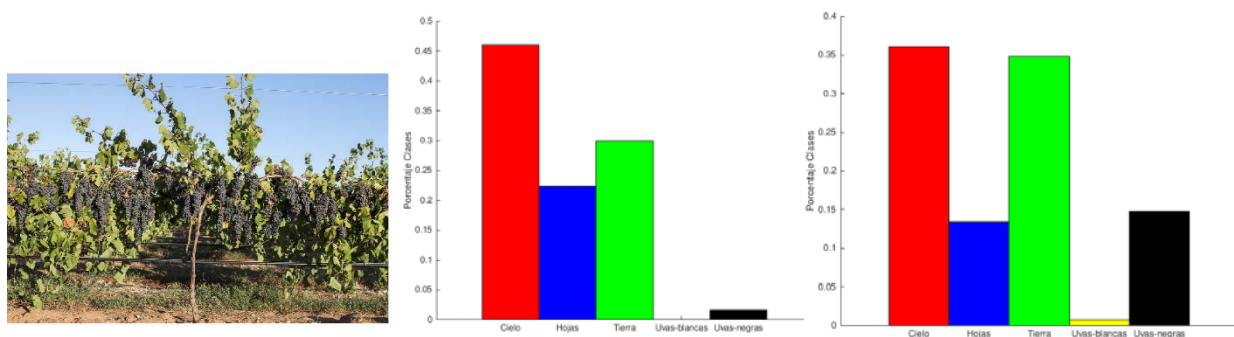


Figura 4.16. Resultados de clasificación para 1200px-Barrosa_Valley_18 con

$W = 200$ y 50 respectivamente

No obstante, conviene reseñar que el incremento del tamaño de la ventana sólo es conveniente para el caso indicado previamente, ya que si esto mismo se aplica sobre la imagen mostrada en la figura 4.13 los resultados empeoran tal y como se muestra en la figura 4.17, donde los porcentajes son: 0.6603 de cielo, 0.0311 de hojas, 0.1115 de tierra, 0.1210 de uvas blancas, 0.0760 de uvas negras. Observándose en este caso el empeoramiento indicado.

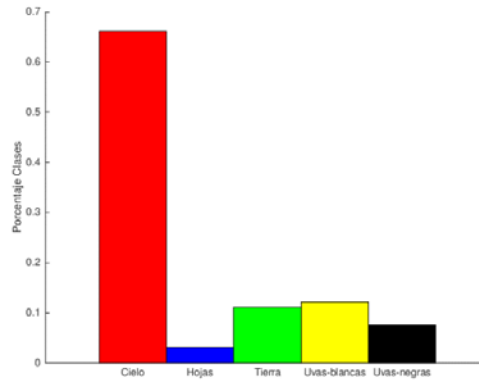


Figura 4.17. Resultados de la red para la imagen viñedo-negras con $w=50$

La mejor solución a este problema consiste en utilizar la misma cámara y configuración óptica, a la vez que el vehículo recorre las líneas de calle de los viñedos de la misma forma y a las mismas distancias (Pajares y col., 2015). De esta forma se obtiene el mismo tipo de imágenes para el entrenamiento de la red y para la clasificación.

En cualquier caso, cabe reseñar, como colofón a la discusión previa, que los resultados observados son aceptables, teniendo en cuenta el relativamente bajo número de imágenes de entrenamiento utilizadas para el aprendizaje, y además con similares tipos de clases y relativa poca diferencia existente entre ellas. Lo ideal, sin duda, es entrenar con el mayor número posible de imágenes, probablemente del orden de miles de imágenes tal y como aparecen los modelos preentrenados y en concreto AlexNet, que conducirá a la obtención de mejores resultados. Si bien, dado que no es éste el objetivo final del trabajo, esta consideración se deja como indicador de posibles mejoras del sistema IoT propuesto.

4.3 Resultados del ajuste mediante regresión

Como se ha explicado en la sección 3.2.2 los datos, para estimar los parámetros del modelo de regresión multivariable para la predicción, se han generado de manera aleatoria dentro de unos parámetros reales de la zona de la Ribera del Duero. El modelo usado para la predicción es el de regresión lineal, explicado en la sección 2.2.1. Los datos almacenados en la nube y utilizados para generar los parámetros del modelo de predicción se encuentran ubicados en la plataforma ThingSpeak, concretamente en el canal nº 690674 accesible a través de la siguiente URL (<https://thingspeak.com/channels/690674>), que se reproducen en la figura 4.18. Por otro

lado, como se observa en la figura 4.19, se dispone de los datos de entrenamiento de las variables dependientes objeto de la predicción, éstos se encuentran en el canal n° 703161 accesible a través de la siguiente URL (<https://thingspeak.com/channels/703161>).

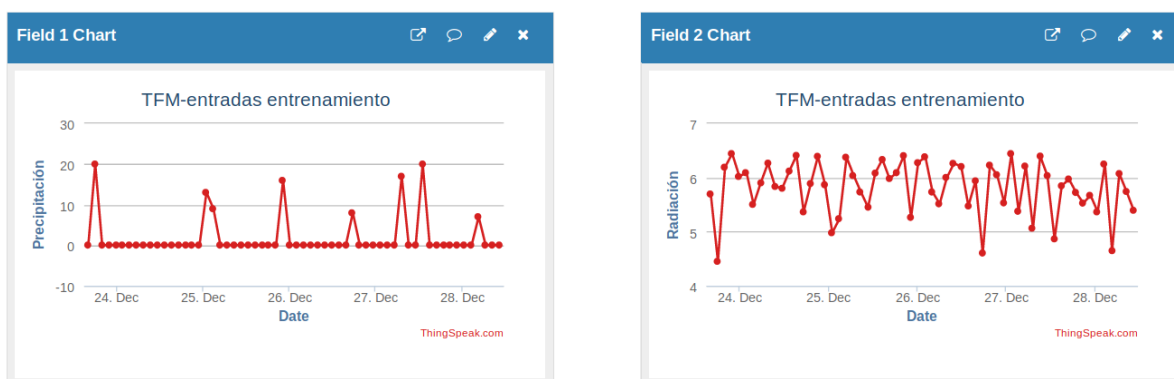


Figura 4.18. Datos de entrada para cálculo de parámetros de regresión linear.

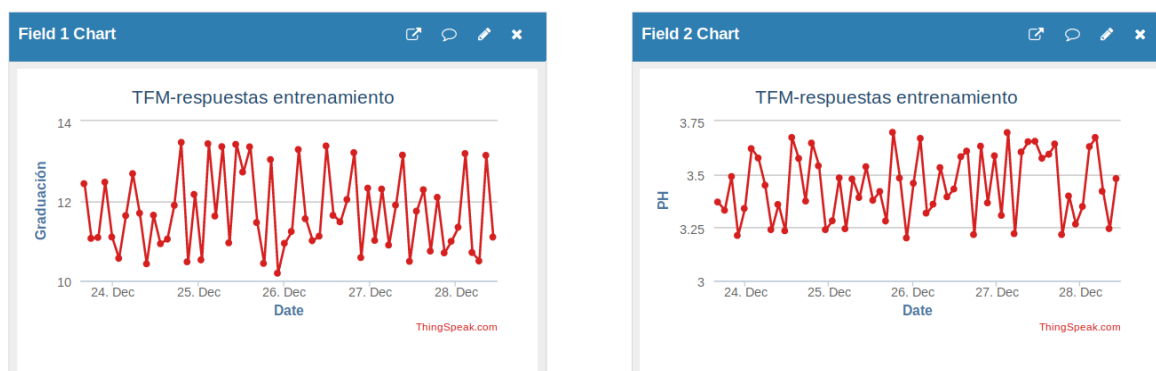


Figura 4.19. Datos de respuesta para cálculo de parámetros de regresión linear

Los parámetros estimados por el modelo diseñado para el predictor mediante la ecuación normal, como se explica en el apartado 2.2.1, se almacenan en el canal n° 748722 accesible a través de la URL (<https://thingspeak.com/channels/748722>). Como se puede observar en la figura 4.20, todos los parámetros obtenidos para las diferentes cinco respuestas se almacenan en un mismo campo del canal, esto es debido a que no se dispone de más canales en la cuenta proporcionada por ThingSpeak debido al limite otorgado por la licencia activa en el momento de realizar los experimentos. Debido a esta limitación en el número de canales disponibles se ha optado por esta solución, si bien lo ideal sería disponer del suficiente número de canales, con sus correspondientes datos asociados, para conseguir un almacenamiento separado.

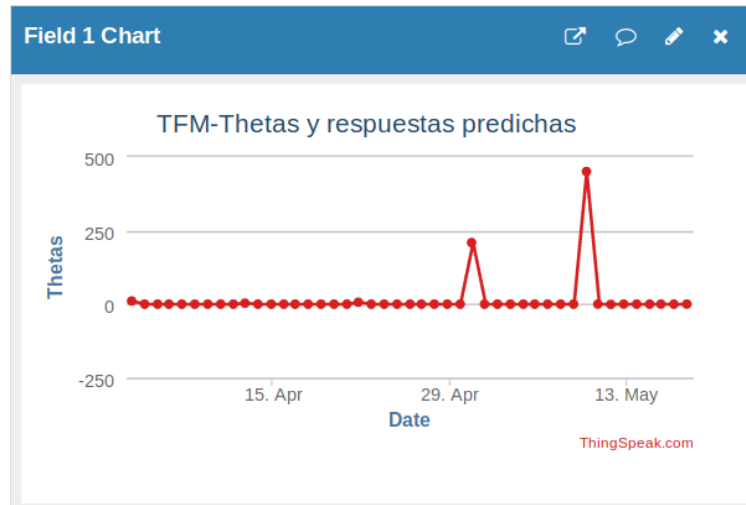


Figura 4.20. Parámetros calculados de la predicción

Por último, se usan los parámetros del canal nº 748722 para el cálculo de variables predichas, para ello se utilizan datos de test ubicados en el canal nº 748742 accesible desde la URL (<https://thingspeak.com/channels/748742>). Finalmente, los datos predichos son subidos al mismo canal que el de los parámetros, esto es el nº 748722.

Capítulo 5 - Conclusiones y trabajo futuro

En el presente trabajo se presenta una solución de concepto en el ámbito de la AP para monitorización de cultivos de viñedos, que demuestra en este caso, cómo mediante el uso de nuevas tecnologías es posible solucionar o paliar los grandes problemas a los que se enfrenta la sociedad de consumo actual. Sin duda la IA y el IoT son paradigmas, que proporcionan tecnologías clave para la sostenibilidad del planeta. Reduciendo el consumo de agua, energía y fertilizantes para optimizar y mejorar la producción de alimento se consigue mejorar no sólo la salud de la Tierra, sino que además permite establecer una sociedad más justa con un reparto más equitativo de su sustento.

Por un lado, el IoT permite determinar en tiempo real la situación del cultivo, así como detectar problemas que se pueden solucionar de manera automática, es precisamente esta velocidad de actuación lo que hace de esta tecnología que los problemas, como puede ser la extensión de plagas, la sequía o la falta de nutrientes en la tierra, no empeoren. Además, con suficiente sensorización es posible localizar con exactitud la ubicación del problema. A modo de ejemplo, la detección de un problema específico como puede ser la falta de agua en determinadas partes del cultivo permitiría abordarlo para actuar sobre ellas de forma localizada. Por otro lado, los predictores ayudan a estimar la producción que va a generar el cultivo en cuestión, así como su calidad en un momento determinado, permitiendo actuaciones concretas para identificar necesidades con el fin de obtener el máximo beneficio. Además, como en toda explotación de cualquier cultivo sólo interesa generar la producción con posibilidades de consumo.

Desde el punto de vista técnico se ha propuesto una solución IoT inteligente a nivel conceptual. Se plantea un despliegue sensorial conectado a través del correspondiente protocolo de comunicaciones y un router de borde, de forma que los datos se envían a un procesador ubicado en un CC, desde el que se suben a la nube, con fines de almacenamiento y para la estimación de los parámetros asociados a un modelo de regresión. Ambos, datos y parámetros estimados se disponen en la nube con carácter público. Dentro de estos datos, existe uno de especial relevancia, que es el cómputo de la razón entre hojas y uvas existentes en una imagen. Para la obtención de este dato, se

plantea la equipación de un sistema robotizado dotado con sendas cámaras de visión que capturan las imágenes lateralmente durante su navegación a lo largo de las calles definidas por las plantas de vides. Las imágenes son enviadas al CC donde se procesan mediante la CNN, que permite detectar incluso rasgos más complejos que los sensores como es la cantidad de fruto que se está produciendo, su coloración y en definitiva la generación de mapas de productividad, así como la posibilidad de detectar enfermedades en los cultivos sin necesidad de intervención humana y con mayor rapidez.

Sin embargo, se han encontrado varios problemas a la hora de realizar este proyecto. Respecto al predictor no existe una manera fiable, por el momento, de que, dadas varias variables de entrada, que definen propiedades del cultivo, predecir varias variables de salida, definiendo propiedades del fruto. Esto supone un problema ya que se requiere crear un modelo de regresión por cada variable dependiente a predecir. Por tanto, se propone como acción de futuro plantear un modelo para resolver la problemática.

Además, dada la imposibilidad de la implantación del sistema en campo, los datos han sido generados, aunque con la coherencia suficiente, de forma aleatoria, por lo que se plantea la evolución hacia la sensorización real.

Tampoco se ha realizado un estudio comparativo de modelos para predecir de manera más exacta las variables dependientes, que, aunque no es un objetivo específico de IoT, sí que resulta de interés, en el ámbito de aplicación del proyecto, por lo que se plantea como un trabajo futuro.

Sobre el análisis de imágenes se ha detectado una deficiencia en el sentido de que no se dispone de una cantidad de imágenes de entrenamiento totalmente acordes con el planteamiento del proyecto. En este trabajo se han utilizado cuarenta por clase con recortes de una misma imagen, en lugar de utilizar miles de imágenes de cada clase procedentes de imágenes captadas en viñedos de forma programada. Los resultados hubiesen sido mucho mejores pudiendo acceder a un banco de imágenes extenso, planteándose como opción de futuro la creación de un banco de datos en esta línea.

Otra idea, que no se ha puesto en práctica es la detección de enfermedades de la uva utilizando CNN mediante una base de datos especialmente dedicada a tal fin. Se trata de una opción de interés para evitar la propagación de enfermedades por la detección temprana.

En cualquier caso, dado el coste inicial del despliegue basado en la solución propuesta, su aplicación se enfocaría hacia grandes extensiones de cultivo, de forma que la solución sea rentable, incluso teniendo en cuenta que algunas soluciones sensoriales resultan económicamente aceptables. En cualquier caso, el objetivo es recuperar la inversión lo más rápido posible, independientemente de los beneficios sociales que la misma pueda ofrecer.

Capítulo 6 - Introduction

6.1 General considerations

The consumption of food, and resources for its cultivation, is increasing every day. United Nations (UN, 2019) in its report concludes that: *"Each year, it is estimated that one third of all food produced, equivalent to 1,300 million tons worth around 1000 million dollars, ends up rotting in the containers of consumers and retailers, or damaged due to malpractices transport and harvesting."* and *"If the world's population were to reach 9.6 billion by 2050, the equivalent of almost three planets would be needed to provide accurate natural resources maintain the current lifestyle"*. Specifically, goal number 12 of the United Nations report foresees to establish by 2030 efficient production means that need less resources to generate only the necessary food required by the planet. From this report it is inferred that the generation of food on a global scale supposes 30% of the energy consumed by the planet and 22% of greenhouse gas production, reason why it is a key objective for the development of humanity.

Not only must we consider the food itself, but also the means and resources required for its production, and one of the great problems of the planet is the lack of fresh water not only for direct human consumption, but also for the animal and irrigation. About 3% of the world's water is suitable for animal consumption and/or cultivation, of which 2.5% is frozen in Antarctica, the Arctic and glaciers. Only 0.5% of the planet's water remains to sustain an ecosystem that is increasingly fragile due to the pollution of these waters.

In the specific case of Spain, the advance of desertification is worrisome (Gobierno de Aragón, 2018). To the water shortages derived from the high consumption and lack of conscience of the society, it is necessary to add the excessive use of water by inefficient crops. According to the report by Rodríguez-Casado et al. (2008), although more than a decade its general conclusions remain valid, the water footprint of Spain stood at a total annual volume of close to 48,000 hm^3 , which meant that in Spain there were 350 reservoirs with a storage capacity out of a total of 54,000 hm^3 of water. Of this consumption more than half of the water to satisfy the demand comes from other countries. Of the total water consumption in Spain more than 80% is used in agriculture. This high-water consumption, coupled with illegal aquifer withdrawals and over-

exploitation of the soil make it less and less cultivable which results, in turn, they need more fertilizer, more energy and raw materials for production, with consequent contamination that these chemicals entail.

In nature everything is connected, and the current level of resource consumption can become unsustainable for the planet Earth in the long term. To solve this problem, or at least mitigate it as much as possible, as has always been done throughout history in different ways, we resort to, and need, the help of new technologies, and this is where the Internet of Things, is called to play an important role. Indeed, thanks to the IoT it is possible to sensorise and interconnect different devices to track crops and means of production on a large scale at very low cost, allowing to carry out detailed studies on what is happening in the development of crops to obtain the final product. Currently, without an interconnected sensorization, it is difficult to determine exactly what is their instantaneous situation, and in real time, as well as to know different parameters that affect them, such as the quality of the earth or humidity among others. And even less how these data affect the final production, in order to be able to consider relevant future aspects depending on the demand of the products, which will undoubtedly translate into cost savings and optimization of resources. With the monitoring of real and updated data in relation to the situation of a certain plantation it is possible to automate the means of production, with human supervision, evolving towards automation, to achieve an efficient performance, based on efficient elaboration processes.

Artificial Intelligence in conjunction with the IoT paradigm offers a wide range of possibilities and solutions for the required and necessary automation, part of which are precisely those presented in this work. Within these solutions, advanced data processing techniques are developed, both locally, and therefore on the side of the sensor in what is called edge computing, as well as remotely. In this work, processing solutions are offered at both levels. Therefore, focusing on the previously mentioned and considering the IoT paradigm as the interconnection of devices (sensory and process), with transfer of data for monitoring and analysis, the work presented focuses on these last two aspects, especially in the analysis, for which intelligent techniques are developed and applied.

Thus, at the local level, the core of automation focuses on the design of a method of image recognition, using deep learning techniques, more specifically, Convolutional

Neural Networks (CNN) to obtain values of a key variable of analysis, relative to determine the degree of development of the crop to include it in the prediction system for production. The variable mentioned is specifically the so-called grapes-leaves ratio. On the other hand, remotely, this is in the cloud, compiles information related to another type of variables (ambient temperature, solar radiation, humidity of the environment and earth, atmospheric pressure), as will be seen later. With all available data estimates are made for prediction purposes, leaving both the data and the estimated parameters available and public use in the corresponding served in the cloud, to perform both subsequent analysis and predictions based on the parameters.

6.2 Motivation

The approach of the concept and analysis of the proposed solution in the present work focuses on the monitoring of a vineyard with denomination of origin. The fact of having chosen this crop and no other types of plantations is based on several reasons, which are summarized below. The first one is that this type of crops generates a national product with great impact within Spanish society and, in addition, there is abundant documentation on these crops, allowing to establish an assessment in this regard. On the other hand, and as a second reason, the vineyard itself needs very specific climatic conditions, among which moderate-low rainfall (400-600 mm as average rainfall per year) or specific soil substrates for its correctness stand out. Development and growth, in addition, and as a third reason, a vineyard to be considered a designation of origin, such as Ribera del Duero, needs to overcome a series of tests related to alcohol content, volatile acidity, total sulphur and reducing sugars, among others. This is achieved by having a very specific grape production marked by the regulation of denomination of origin. Cultivation of grapes cannot be exceeded, because the properties of the same would not comply with current regulations. On the other hand, low productions would cause undesirable economic losses. All these circumstances motivate the need for monitoring over time, while the data is available for use by the community concerned and the corresponding analyses are carried out to establish corrective measures, where the IoT provides a feasible solution. To this problem.

A vineyard of denomination of origin like that of Ribera del Duero is a good object of study, not only to be able to extrapolate the work done to other more resistant crops

such as cereals or corn that do not need environments with such regularization, but also, because it needs to have a quantity of concrete production without surpluses or shortages, thus fulfilling the objective number twelve of the United Nations (UN, 2019).

From the technical point of view, the motivation of the work is sufficiently justified. Indeed, the concept of development proposed under the IoT paradigm is based on the capture of data from sensory sources, which, as previously stated, are processed partly locally and partly in the cloud. Given that there is no real implantation of devices at the field level, the capture is carried out, in the context of the IoT, in a simulated manner. Thus, enough data are available to determine the variables previously analysed. Therefore, the approach offers a conceptual and viable solution in the IoT environment.

6.3 Objectives

The general objective that is proposed in this paper is to design an IoT system, as a possible and feasible solution to analyse and optimize, through the application of intelligent techniques, a vineyard crop to fulfil the United Nations objective number twelve to guarantee modalities of sustainable consumption and production, as previously mentioned. Specifically, it is a solution applicable to a vineyard to produce wine with designation of origin Ribera del Duero. Since the implantation of sensory technologies in situ is not possible, the approach proposes the capture of data from repositories available on a public basis, including the use of images for processing. In this way, a generic solution is guaranteed, so that for its real implementation, there is no more to replace the sources with the corresponding sensors that generate the same type of data.

Within the specific objectives and from the global point of view of the IoT, it is proposed to develop an intelligent system through specific techniques of deep learning, more specifically of the CNN type. This network processes the available images in local mode to obtain the percentage of vineyard leaves in relation to the amount of grapes present. This processed data is uploaded and published remotely (cloud), being available for public visualization and subsequent processing also in remote mode, together with the rest of available data (ambient temperature, solar radiation, humidity of the environment and earth, atmospheric pressure) with prediction purposes. For the prediction multiple linear regression techniques are used. The data, in the absence of

real sensory data of a crop, are generated in a random manner based on real data from the AEMET (2019) of the meteorological station of Aranda de Duero (Burgos).

Therefore, the specific objectives to achieve the integrated IoT solution are summarized as follows:

1. Design of a procedure to obtain data.
2. Development and integration, at the local level, of an image processing method based on CNN, which includes two training and classification modules.
3. Design and configuration of the remote platform for storage, processing and publication of data in the cloud.
4. Analysis of the data using multivariate regression techniques, obtaining the corresponding parameters of the model, which are published in the cloud.
5. Integration of the modules that implement the previous processes.
6. Analysis and evaluation of the results.

6.4 Content of this document

The report is structured in five clearly differentiated chapters detailing the research carried out and the development of the project.

The first chapter details the general approach of the project, specifying as general aspects the reasons that motivate the proposed IoT solution, to solve the problem of crop data monitoring. In addition, the objectives proposed in this work are included.

In the second chapter, a general review is made on IoT applications in agriculture, including precision, in general and in vineyards. It also explains the theoretical foundations of the CNN, as well as the regression techniques. These are the data processing techniques used both locally (CNN) and remotely (regression).

In the third chapter the proposed architectural solution is exposed, including the local and remote modules, including the established communication mechanisms, as well as the publication of the data available in the cloud.

The fourth chapter provides the results obtained in relation to the different modules developed, as well as those related to the analysis of the processes carried out.

In the fifth chapter the derived conclusions are included, as well as the possible future works.

Finally, the complementary annexes are included as the code used in the application.

Capítulo 7 - Conclusions and future work

This report presents a concept solution in the field of PA for monitoring vineyard crops, which demonstrates in this case, how using new technologies it is possible to solve or alleviate the major problems facing the society of current consumption. Definitely, the AI and the IoT are paradigms, which provide key technologies for the sustainability of the planet. Reducing the consumption of water, energy and fertilizers to optimize and improve the production of food can improve not only the health of the Earth, but also allows to establish a more just society with a more equitable distribution of their livelihood.

On the one hand, the IoT allows detecting in real time the situation of the crop, as well as detecting problems that can be solved automatically, it is precisely this speed of action that makes this technology that the problems, such as the extension of pests, drought or lack of nutrients in the land, do not get worse. In addition, with enough sensorization it is possible to determinate the exact location of the problem. As an example, the detection of a specific problem such as the lack of water in certain parts of the crop would allow it to be addressed in order to act on them in a localized manner. On the other hand, predictors help to estimate the production that the crop in question will produce, as well as its quality at a specific time, allowing specific actions to identify needs to obtain the maximum benefit. In addition, as in any exploitation of any crop, it is only interesting to generate production with consumption possibilities. From the technical point of view, an intelligent IoT solution has been proposed at a conceptual level. A sensory deployment connected through the corresponding communications protocol and an edge router is proposed, so that the data is sent to a processor located in a CC, from which they are uploaded to the cloud, for storage purposes and for the estimation of the parameters associated with a regression model. Both, data and estimated parameters are available in the cloud with public character. Within these data, there is one of special relevance, which is the computation of the ratio between leaves and grapes in an image. To obtain this information, the equipment of a robotic system equipped with two vision cameras that capture the images laterally while navigating along the streets defined by the vines plants is considered. The images are sent to the CC where they are processed through CNN, which allows to detect even more complex features than the sensors such as the amount of fruit that is being

produced, its coloration and ultimately the generation of productivity maps, as well as the possibility to detect diseases in crops without the need for human intervention and more quickly.

However, several problems have been found during the development of this project. Regarding the predictor, there is no reliable way, now, that given several input variables, which define crop properties, allow to several output variables, defining properties of the fruit. This is a problem since it is required to create a regression model for each dependent variable to predict. Therefore, it is proposed as a future action to propose a model to solve the problem.

In addition, the data have been generated, even though with some coherence, in a random manner, so the evolution towards real sensing is considered. A comparative study of models has not been carried out to more accurately predict the dependent variables, which, although it is not a specific IoT objective, is of interest in the scope of the project, so it is proposed as a future task.

On the analysis of images, a deficiency has been detected in the sense that there is not a quantity of training images fully consistent with the project's approach. In this work forty have been used per class with cuts of the same image, instead of using thousands of images of each class from images captured in vineyards in a programmed manner. The results would have been much better, being able to access an extensive image bank, considering the creation of a databank in this line as a future option.

Another idea that has not been put into practice is the detection of grape diseases by CNN through a database specially dedicated to this purpose. It is an option of interest to avoid the spread of diseases by early detection.

In any case, given the initial cost of deployment based on the proposed solution, its application would focus on large areas of crops, so that the solution is useful, even considering that some sensory solutions are economically acceptable. In any case, the objective is to recover the investment as quickly as possible, regardless of the social benefits that it may offer.

Bibliografía

1. AEMET (2019). Documentos de la AEMET. *Atlas Agroclimático de Castilla y León*. Disponible en línea: https://www.aemet.es/documentos/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/Atlas_Agroclimatico_CyL2013_00/Atlas_Agroclimatico_CyL2013_00.pdf (accedido Febrero 2019).
2. Ashton, K. (2009) *That 'Internet of Things' Thing*. RFID journal. (2009). Disponible en línea: <http://www.itrc.jp/libraries/RFIDjournal-That%20Internet%20of%20Things%20Thing.pdf> (accedido Febrero 2019).
3. Burgos-Artizzu, X.P., Ribeiro, A., Guijarro, M., Pajares, G. (2011). *Real-time image processing for crop/weed discrimination in maize fields*. Computers and Electronics in Agriculture, 75(2), 337-346.
4. Charniak, E., McDermott, D. (1985). *Introduction to Artificial Intelligence*. Addison-Wesley. Readings, MA.
5. Daugman, J.G. (1980). *Two-dimensional spectral analysis of cortical receptive field profiles*. Vision Res., 20(10), 847-56.
6. Daugman, J.G. (1985). *Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters*. Journal of the Optical Society of America A, 2, 1160-1169.
7. Duda, R.O., Hart, P.E., Stork, D.G. (2001). *Pattern Classification*. John Wiley & Sons, NY.
8. Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT Press. Disponible en línea: <https://www.deeplearningbook.org> (accedido febrero 2019).
9. Guerrero, J.M., Ruz, J.J., Pajares, G. (2017). *Crop rows and weeds detection in maize fields applying a computer vision system based on geometry*. Computers and Electronics in Agriculture, 142 (Part A), 461-472.
10. GoogleLeNet (2019). *BVLC GoogLeNet Model*. Disponible en línea: https://github.com/BVLC/caffe/tree/master/models/bvlc_googlenet. (accedido Febrero 2019).
11. Haugeland, J. (1985). *AI: the very idea*. MIT Press, Cambridge, MA.
12. iQonsulting (2008) *Tecnologías Aplicables a la Agricultura de Precisión* (García, M.J. coordinadora) Plataforma silvoagropecuaria, Gobierno de Chile. Disponible en línea: http://wapp.corfo.cl/ticrural/descargas/I+D+I_3_TECNOLOGIAS_AGRICULTURA_PR

- [ECISION.pdf](#) (accedido Marzo 2019).
13. Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. In Proc. 25th Int. Conf. on Neural Information Processing Systems (NIPS'12), vol. 1, pp. 1097-1105.
 14. Libelium (2019), *Agriculture v30 Board*. Disponible en línea: http://www.libelium.com/downloads/documentation/agriculture_sensor_board_3.0.pdf (accedido Febrero 2019).
 15. Mathworks (2019). MATLAB. Disponible en línea: <https://es.mathworks.com/products/matlab.html> (accedido Febrero 2019).
 16. Navas, L.M., González, F., Gómez-Gil, J., Ruiz-Ruiz, G., Barreiro, P., Andújar, D., Diezma, B., Báguena, E., Garrido, M. (Valero, C. coordinador) (2010). *Ahorro y eficiencia energética en la agricultura*. Instituto para la Diversificación y Ahorro de la Energía (IDAE). Disponible en línea: http://www.idae.es/uploads/documentos/documentos_10995_AEE_Agr_Precision_AGR16_A2010_69054ba0.pdf (accedido Marzo 2019).
 17. Pajares, G.; García-Santillán, I.; Campos, Y.; Montalvo, M.; Guerrero, J.M.; Emmi, L.; Romeo, J.; Guijarro, M.; Gonzalez-de-Santos, P. (2016). *Machine-Vision Systems Selection for Agricultural Vehicles: A Guide*. J. Imaging, 2, 34.
 18. PIM Ribera Duero (2018). *Parcelas Índices de Maduración*. Consejo Regulador del Ribera del Duero (2018). Disponible en línea: https://riberadelduero.es/sites/default/files/viticultura/indices_de_maduracion_08_de_octubre_de_2018_0.pdf (accedido Marzo 2019).
 19. PIXNIO (2019a) Imágenes de Viñedos (Licencia CC0). Disponible en línea: <https://pixnio.com/es/plantas/frutas/uvas/hoja-agricultura-frutales-vid-vina-naturaleza> (accedido Febrero 2019).
 20. PIXNIO (2019b) Imágenes de Viñedos (Licencia CC0). Disponible en línea: <https://pixnio.com/es/plantas/frutas/uvas/viticultura-naturaleza-vid-vinedo-hoja-fruta-agricultura> (accedido Febrero 2019).
 21. PIXNIO (2019c). uva-negra. (Licencia CC0). Disponible en línea: <https://pixnio.com/es/plantas/frutas/uvas/alimento-hoja-vinedo-agricultura-fruta-naturaleza-uva-uvass> (accedido Febrero 2019).
 22. PIXNIO (2019d). vinedo-mixto2. (Licencia CC0). Disponible en línea: <https://pixnio.com/es/plantas/frutas/uvas/agricultura-uva-grapevine-hoja-planta-vina-uvass-fruta> (accedido Febrero 2019).

23. Poole, D. L., Mackworth, A. K., Goebel, R. (1998). *Computational intelligence: a logical approach* (Vol. 1). Oxford University Press, NY.
24. ResNet-18 (2019). He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.
25. Ribera del Duero (2018). *Normativa para clasificación denominación de origen del Ribera del Duero*. Consejo de denominación de origen de Ribera del Duero. Disponible en línea: https://riberadelduero.es/sites/default/files/consejo-regulador/normativa_calificacion_vinos_sin_c_i.pdf (accedido Febrero 2019).
26. Rich, E., Knight. K. (1991) *Inteligência Artificial*. McGraw-Hill, NY.
27. RHEA (2014). *Robot Fleets for Highly Effective Agriculture and Forestry Management* (NMP-CP-IP 245986-2 RHEA). Disponible en línea: <http://www.rhea-project.eu/> (accedido Marzo 2019).
28. Rodríguez-Casado, R., Garrido-Colmenero, A., Llamas-Madurga, M.R., Varela-Ortega, C. (2008). *La huella hidrológica de la Agricultura española*. Disponible en línea: <http://oa.upm.es/2146/> (accedido Enero 2019).
29. Russakovsky, O., Deng, J., Su, H. Krause, J., Satheesh, S., Ma, S. Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L. (2015). *ImageNet Large Scale Visual Recognition Challenge*. Int. J. Computer Vision (IJCV). 115(3), 211–252.
30. Sancho, A. (2019), Libelium. *Webinar Smart Vineyards*. Disponible en línea: <http://www.libelium.com/free-live-iot-overview-webinars/> (accedido Marzo 2019).
31. Simonyan, K., Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. arXiv preprint arXiv:1409.1556.
32. Tellaeche, A., Pajares, G., Burgos-Artizzu, X.P., Ribeiro, A. (2011). *A computer vision approach for weeds identification through Support Vector Machines*. Applied Soft Computing 11 (1), 908-915.
33. ThingSpeak (2019). Understand Your Things The open IoT platform with MATLAB analytics. Disponible en línea: <https://thingspeak.com/> (accedido Febrero 2019).
34. UN (2019) Naciones Unidas, *Objetivo 12: Garantizar modalidades de consumo y producción sostenibles*. Disponible en línea: <https://www.un.org/sustainabledevelopment/es/sustainable-consumption-production/> (accedido Abril, 2019).

35. Unplash (2019a). Marius-ciocirlan-414588-unsplash. (Licencia CC). Disponible en línea: <https://unsplash.com/photos/Llqq7lv-Rjk> (accedido Febrero 2019).
36. Unplash (2019b). Samuel-zeller-339493-unsplash. (Licencia CC). Disponible en línea: <https://unsplash.com/photos/piR6qwg11Fo> (accedido Febrero 2019).
37. VineScout (2019). VineScout: *Intelligent decisions for Vineyards robots*. Disponible en línea: <http://vinescout.eu/web/> (accedido Febrero 2019).
38. VGG-16, VGG-19 (2019). *Very Deep Convolutional Networks for Large-Scale Visual Recognition*. Disponible en línea: http://www.robots.ox.ac.uk/~vgg/research/very_deep/ (accedido Febrero 2019).
39. Wikimedia (2019a). Grapes. (Licencia CC0). Disponible en línea: https://commons.wikimedia.org/wiki/File:Grapes_1.jpg (accedido Febrero 2019).
40. Wikimedia (2019b). Montepulciano_d_Abruzzo_04_(RaBoe) (Licencia CC). Disponible en línea: [https://commons.wikimedia.org/wiki/File:Montepulciano_d_Abruzzo_04_\(RaBoe\).jpg](https://commons.wikimedia.org/wiki/File:Montepulciano_d_Abruzzo_04_(RaBoe).jpg) (accedido Febrero 2019).
41. Wikimedia (2019c). Tannat_Aout_1. (Licencia Free Art Licence). Disponible en línea: https://upload.wikimedia.org/wikipedia/commons/b/b4/Tannat_Aout_1.jpg (accedido Febrero 2019).
42. Wikimedia (2019d). Trte. (Licencia CC) Disponible en línea: <https://commons.wikimedia.org/wiki/File:Trte.jpg> (accedido Febrero 2019).
43. Wikimedia (2019e). Cepa Vieja de Mencía. (Licencia CC) Disponible en línea: https://commons.wikimedia.org/wiki/File:Cepa_Vieja_de_Menc%C3%ADA.jpg (accedido Febrero 2019).
44. Wikimedia (2019f). Barossa Valley plantation (Licencia CC). Disponible en línea: https://commons.wikimedia.org/wiki/File:Barossa_Valley_plantation.jpg (accedido Febrero 2019).
45. Wikimedia (2019g). 1024px-Iberieli_Wineyards (Licencia CC). Disponible en línea: https://commons.wikimedia.org/wiki/File:Iberieli_Wineyards.jpg (accedido Febrero 2019).
46. Wikimedia (2019h). Barossa_Valley_14. (Licencia CC) Disponible en línea: https://commons.wikimedia.org/wiki/File:Barossa_Valley_14.jpg (accedido Febrero 2019).

47. Wikimedia (2019i). Galician_vineyard_with_wide_vine_spacing. (Licencia CC). Disponible en línea: https://commons.wikimedia.org/wiki/File:Galician_vineyard_with_wide_vine_spacing.jpg (accedido Febrero 2019).
48. Wikimedia (2019j). Growing-grapes. (Licencia CC0). Disponible en línea: <https://commons.wikimedia.org/wiki/File:Growing-grapes.jpg> (accedido Febrero 2019).
49. Wikimedia (2019k). Australia_Red_Grapes. (Licencia CC0). Disponible en línea: [https://commons.wikimedia.org/wiki/File:Australia_Red_Grapes_\(3406840850\).jpg](https://commons.wikimedia.org/wiki/File:Australia_Red_Grapes_(3406840850).jpg) (accedido Febrero 2019).
50. Wikimedia (2019l). Semillon grapes at Château Doisy Védrières, Barsac, Sauternes (Licencia CC) Disponible en línea: https://upload.wikimedia.org/wikipedia/commons/b/b4/Semillon_grapes_at_Ch%C3%A2teau_Doisy-V%C3%A9drines%2C_Barsac%2C_Sauternes.jpg (accedido Febrero 2019).

Anexo: Datos y código

Junto con la memoria se realiza la entrega de los datos necesarios para la ejecución del código y los programas correspondientes.

Datos

- a) Directorio con las imágenes de entrenamiento de la CNN (imágenes).
- b) Directorio con las imágenes de prueba para analizar la clasificación de éstas por la CNN (imágenes completas).
- c) Ficheros en formato csv con los datos generados con las variables de análisis del viñedo (data_entradas.csv y data_respuestas.csv).

Código

- a) Programa de entrenamiento de la CNN en lenguaje MATLAB (CNNTraning.m).
- b) Programa de clasificación de imágenes en lenguaje MATLAB (CNNClasificacionUvas.m).
- c) Generador de datos aleatorios en lenguaje Python (csv_generator.py).
- d) Programa de estimador del regresor lineal (Estimador.m).
- e) Programa de predicción del regresor lineal (Predictor.m).

Ejecución de los programas

- 1) Crear una cuenta en ThingSpeak con los siguientes canales.
 - a. TFM-entradas entrenamiento. Con los siguientes campos:

Field 1	Precipitación	<input checked="" type="checkbox"/>
Field 2	Radiación	<input checked="" type="checkbox"/>
Field 3	Temperatura Atm	<input checked="" type="checkbox"/>
Field 4	Humedad Atm	<input checked="" type="checkbox"/>
Field 5	Presión Atm	<input checked="" type="checkbox"/>
Field 6	Temperatura Suelo	<input checked="" type="checkbox"/>
Field 7	Humedad Suelo	<input checked="" type="checkbox"/>
Field 8	Diferencia uva/hoja	<input checked="" type="checkbox"/>

b. TFM-respuestas entrenamiento. Con los siguientes campos:

Field 1	Graduación	<input checked="" type="checkbox"/>
Field 2	PH	<input checked="" type="checkbox"/>
Field 3	Acidez	<input checked="" type="checkbox"/>
Field 4	Peso	<input checked="" type="checkbox"/>
Field 5	Antocitos	<input checked="" type="checkbox"/>

c. TFM-Thetas y respuestas predichas. Con los siguientes campos:

Field 1	Thetas	<input checked="" type="checkbox"/>
Field 2	Graduación	<input checked="" type="checkbox"/>
Field 3	PH	<input checked="" type="checkbox"/>
Field 4	Acidez	<input checked="" type="checkbox"/>
Field 5	Peso	<input checked="" type="checkbox"/>
Field 6	Antocitos	<input checked="" type="checkbox"/>

d. TFM-entradas test. Con los siguientes campos:

Field 1	Precipitación	<input checked="" type="checkbox"/>
Field 2	Radiación	<input checked="" type="checkbox"/>
Field 3	Temperatura Atm	<input checked="" type="checkbox"/>
Field 4	Humedad Atm	<input checked="" type="checkbox"/>
Field 5	Presión Atm	<input checked="" type="checkbox"/>
Field 6	Temperatura Suelo	<input checked="" type="checkbox"/>
Field 7	Humedad Suelo	<input checked="" type="checkbox"/>
Field 8	Diferencia uva/hoja	<input checked="" type="checkbox"/>

2) Entrenamiento de la red ejecutando el programa CNNTTraining.m. Éste generará los parámetros aprendidos por la red almacenándolos en el

fichero NetTransferUvas.mat. Este programa tiene que tener definido el directorio donde estén las imágenes de entrenamiento (imágenes).

- 3) Ejecutar el programa CNNClasificacionUvas.m seleccionando una imagen del directorio "imágenes completas" éste subirá su resultado al campo 8 del canal "TFM-entradas test".
- 4) Ejecutar el generador de datos aleatorios (csv_generator.py). Éste generará los ficheros data_entradas.csv y data_respuestas.csv. Estos datos se subirán al ThingSpeak a los canales "TFM-entradas entrenamiento" y "TFM-respuestas entrenamiento" respectivamente.
- 5) Ejecutar Estimador.m, cargado previamente en ThingSpeak para su ejecución en la nube. Éste generará los parámetros del modelo de regresión y los almacenará en el campo 1 del canal "TFM-Thetas y respuestas predichas".
- 6) Ejecutar Predictor.m, cargado previamente en ThingSpeak para su ejecución en la nube. Éste cargará las estimaciones de las propiedades de la uva en el canal "TFM-Thetas y respuestas predichas" en función de los datos del canal "TFM-entradas test".